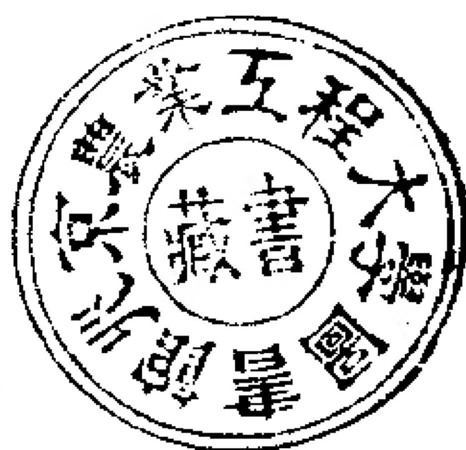
时间序列分析与动态数据建模

杨位钦　顾　岚　编著



北京工业学院出版社

内 容 简 介

本书首先讨论采样数据的检验和预处理，然后介绍平稳随机过程的基本特征及其模型，以线性 ARMA 模型为主，随后两个重点部分是时间域模型和频率域模型的估计方法，包括极大熵谱估计。在建模的基础上介绍了几种时间序列的预报方法。除单变量模型外，本书也讨论多变量的时域和频域模型及其估计，对趋势性、季节性、混合回归、疏系数、门限回归、自适应、双线性、指数等特定形式的模型也做了不同程度的讨论，各部分都附有可在IBM-PC/XT微机上运行的Fortran程序及各种实例分析。本书可以作为大学生、研究生的教材和有关科技人员的参考书。

本书经兵器工业部第一教材编审委员会自动控制编审小组审定并推荐作为部统编教材。

时间序列分析与动态数据建模

杨位钦　顾　岚编著

•

北京工业学院出版社出版

新华书店北京发行所发行　各地新华书店经售

通县向阳印刷厂印刷

•

850×1168毫米　32开本　21.25印张530千字

1986年12月第一版　1986年12月第一次印刷

印数：1—4,000册

统一书号：15434·36　定价：3.55元

前　言

对生产和科学研究等过程中的某一个变量或一组变量$x(t)$进行观察测量，在一系列时刻$t_1, t_2 \cdots\cdots t_n$（$t$为自变量且$t_1 < t_2 < \cdots < t_N$）得到的离散的有序数集合$x(t_1)$，$x(t_2)$，……，$x(t_i)$，……，$x(t_N)$称为离散数字时间序列（本书只讨论离散情况，简称为时间序列）。在实际问题中，自变量t除代表时间外，也可具有各种不同的物理意义，例如代表长度、温度、速度或其它单调递增地取值的物理量。

气象上的月降水量序列，水文上的河流流量序列，海洋上的海浪序列，天文上的太阳黑子序列，地球物理上的地震波序列，医学上的脑电波序列，机械系统的振动序列，雷达系统对活动目标的跟踪量测序列，化工反应系统的输入输出函数序列，自然界某种生物总数的消长序列，商业经济方面的产值、销售额的序列等等都是时间序列的具体例子。

用来分析各种相依有序的离散数据集合的整个方法称为时间序列分析。一般说来，时间序列很难用一个完全确定的函数或函数组给出来，它们大都具有统计规律性，可以通过概率分布函数或函数组对它们取值的规律性作统计描述，这种类型的时间序列称为随机的。本书主要讨论时间序列的统计分析和模型构造问题。

在生产、科研、社会生活和国防建设中，时间序列分析有着广泛的实际应用，其中主要包括：

1．分析时间序列的统计规律性，推断产生时间序列的物理系统的性质，找出它的物理规律。

2．根据对时间序列统计规律性的分析，构造拟合它的最佳

数学模型，浓缩时间序列的信息，简化对时间序列的表示。

3．利用拟合的数学模型预报时间序列未来的可能取值，给出预报结果的精度分析。

4．根据拟合的数学模型模拟出时间序列，用于分析和优化处理。

时间序列分析属于统计学科的分支，近十多年来的发展很活跃，原有的方法有很多新进展，本书在介绍时序分析的基本原理和方法的时候注意具体应用，为便于读者进行时序分析和建模的实践，在附录中提供了能在IBM-PC/XT微机上使用的FORTRAN程序，书中的计算例子除某些引用的结果外都是按照这些程序进行处理的。应当说明，虽然提供的程序具有一定的实用性，但在编写上并非都是最优的。

本书的编写得到了北京工业学院张志方教授的热情鼓励和支持，并承蒙中国科学院应用数学所陈兆国副研究员主审，提出了宝贵的意见，我们在此表示深切的谢意。

编者

一九八六年三月

目　录

第一章　采样数据的检验和预处理

在实际问题中，测量记录常常是时间的连续函数，为了数字计算处理上的方便，往往只按一定的时间间隔取值，怎样合理地选择采样间隔，使得采样所得的序列能如实地代表原来的时间序列，这是数据采集首先遇到的一个重要问题。

在得到大量的量测数据之后，还要进行初步的整理和必要的检验，以拗去粗取精，为进一步深入分析提供较好的依据和参考。这些工作都属于预处理。

1.1　采样间隔和频率混叠

设记录到的连续时间t的函数为$x(t)$（$t_0 \leqslant t \leqslant t_0+T$），这里只讨论在等间隔上取值的时间序列

$$x_1,\ x_2,\ \cdots,\ x_i,\ \cdots,\ x_N \tag{1-1-1}$$

其中$x_i=x(t_i)=x(t_0+i\Delta)$。

在时间序列$x(t)$的采样过程中，取不同的采样间隔Δ可以得到不同的数字时间序列。这种离散的序列抛弃了$x(t)$在$t_0+(i-1)\Delta$和$t_0+i\Delta$之间的值，即这部分信息在采样后丧失了。显然，Δ愈小，(1-1-1)和$x(t)$的差别就愈小，信息损失就愈小。但随着Δ的减小，采样和统计处理的数据量$N=T/\Delta$也就愈大，需要的计算机存储单元和计算时间（如建模所需的时间）也就增加了，所以采样间隔不能无限制地减小。

申农（*Shannon*）采样定理（详见附录一2）给出了理论的采样时间间隔，规定了能从采样数据复现原来信号所必须的最小

采样频率（或最大采样间隔）。尽管由于很难事先确知时间序列的总体频谱特性，我们还是可以根据经验、直观和实际的物理要求，参照采样定理，在不过分减小信息损失和不过分增加数据量之间作出合理的选择。

在反映时间序列的原有信息方面，有一个频率混叠现象是和Δ有直接关系的。问题的提出是：在给定的Δ下，离散序列 x_t 的频率分析能不能得出$x(t)$原来具有的频率分量？

以对连续的谐波信号进行等间隔采样为例，可以说明频率混叠现象。图1-1中的实线为原来的连续信号，而虚线是根据采样点连接起来的。

从这个图可以归纳出离散化采样处理所产生的数字时间序列的某些性质：

（1）如果实际信号频率小于或等于 $1/(2\Delta)$Hz 即采样频率之半，则由采样点反映的频率和实际信号频率相同。（见图1-1 a，b，c，d）。

（2）如果实际信号频率较高（图1-1 *e*，*f*，*g*，*h*，*i*），则根据采样点来“重建”原来信号时，就会得到比实际频率低的波形（在*f*和*i*的情况下，甚至得到的不是正弦波而是直流信号）。

（3）采样间隔Δ确定了频率分析中的最高分辨频率。直观地看，要确定出一个周期所需要的最少信息是三个采样点，即2Δ的时间区间，也就是说，以Δ为间隔采样时，能够检测出的最短周期为2Δ。或者说，能够分辨的最高频率是$1/(2\Delta)$。这个极限频率称之为乃奎斯特(*Nyquist*)频率。它也等于采样频率$f_s=1/\Delta$的一半。

上述性质可以从正弦信号的周期性得出，设

f_c=连续信号的真实频率

f_a=离散化信号的频率

则离散化的采样信号$\cos 2\pi f_a t$可表为

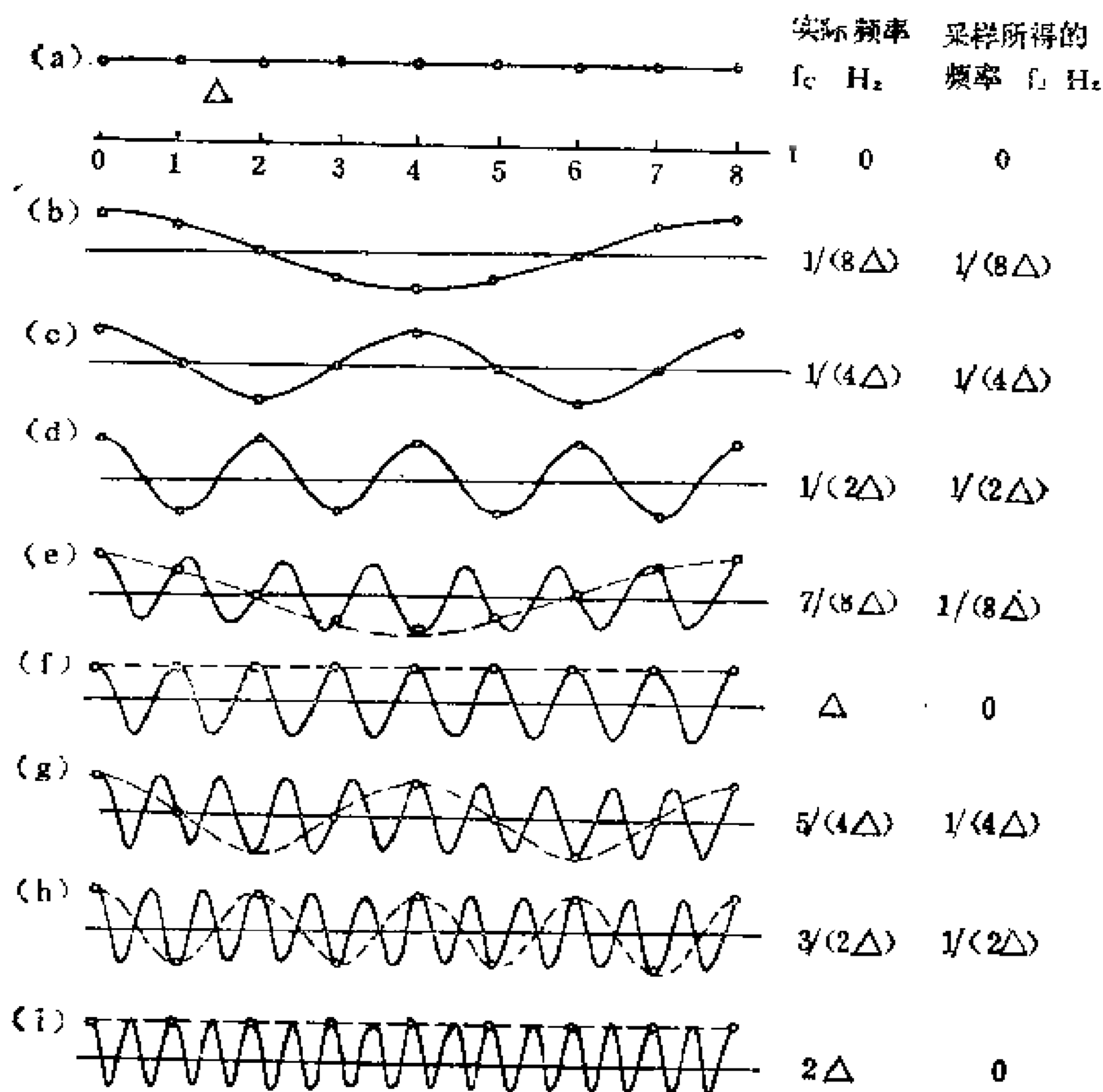


图1-1 正弦信号的频率混叠现象

$$
\begin{aligned}
\cos 2\pi f_C t\big|_{t=K} &= \cos 2\pi f_a k\Delta \\
&= \cos 2\pi(1/\Delta \pm f_a)k\Delta \\
&= \cos 2\pi(2/\Delta \pm f_a)k\Delta \qquad\qquad (1\text{-}1\text{-}2) \\
&= \cos 2\pi(3/\Delta \pm f_a)k\Delta \qquad k=0,\ 1,\ 2,\ \cdots \\
&\ \ \vdots
\end{aligned}
$$

可见，对于同样的采样信号频率(f_a)，可以有许许多多不同频率(f_e)的连续信号与之对应。例如若$f_a=\frac{1}{8\Delta}$，则可有

$$
f_e=\begin{cases}1/\Delta\pm f_a=\begin{cases}1/\Delta-1/(8\Delta)=7/(8\Delta)\\1/\Delta+1/(8\Delta)=9/(8\Delta)\end{cases}\\2/\Delta\pm f_a=\begin{cases}2/\Delta-1/(8\Delta)=15/(8\Delta)\\2/\Delta+1/(8\Delta)=17/(8\Delta)\end{cases}\\\quad\vdots\end{cases}\qquad(1\text{-}1\text{-}3)
$$

这意味着，如果采样得到的信号频率为$f_a=1/(8\Delta)$，我们就难于断定实际信号的频率究竟是$7/(8\Delta)$，还是$9/(8\Delta)$，还是$15/(8\Delta)\cdots$，反过来说，如果实际信号具有$f_e=1/(8\Delta),7/(8\Delta)$等多种频率成份，那么这些频率成份对于构成离散信号中频率为$1/(8\Delta)$的分量都有贡献。即便实际信号并不存在$1/(8\Delta)$的频率，采样信号仍可有该频率成分存在，因为连续信号中的$7/(8\Delta)$和$9/(8\Delta)$，$15/(8\Delta)\cdots$的频率成分也会在$f_a=1/(8\Delta)$的采样信号中反映出来。这种在采样数据中不希望有的效应就是"频率混叠"。

$$
f_a\Longleftrightarrow f_e=n/\Delta\pm f_a\qquad n=0,\ 1,\ 2,\ \cdots\qquad(1\text{-}1\text{-}4)
$$

图1-2表示出f_a和f_e之间的多值对应关系。可以看出，各个f_e均对称于$1/(2\Delta)$，$1/\Delta,3/(2\Delta),2/\Delta,\cdots$（即$f_{max},2f_{max},3f_{max},\cdots$），所以这些频率也称折叠频率。

前已指出，靠无限减小Δ（即增大折叠频率）是不现实的，它也会使出现频率混叠的范围扩大。通常是利用低通滤波处理的方法来解决频率混叠效应。假设低通滤波器是理想的(图1-3a)，则在$f>f_c$以外的频率成分被滤掉了，因此不会在f_c范围内引起混叠。不过，实际滤波器频率响应的模是逐渐衰减的（图1-3b），f_c以外的信号并没有完全滤掉。通常截止频率f_c定义在$M=0.707$

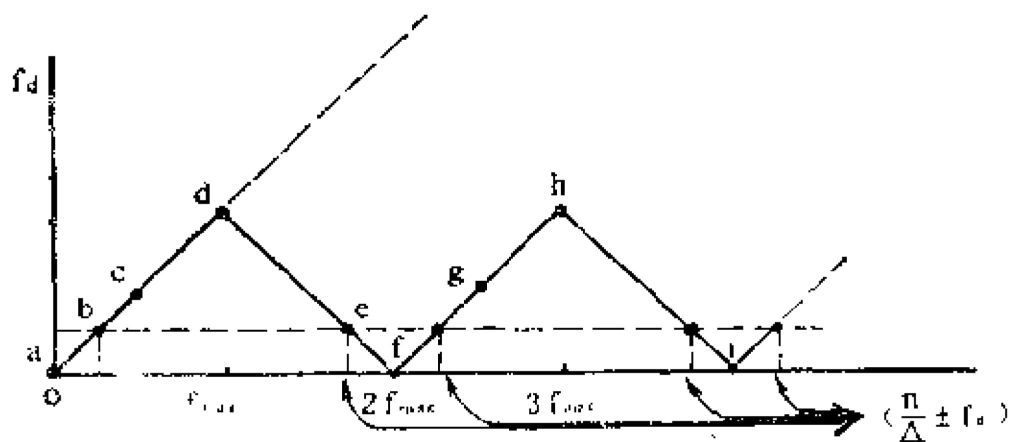


图1-2 连续信号频率与采样信号频率的关系

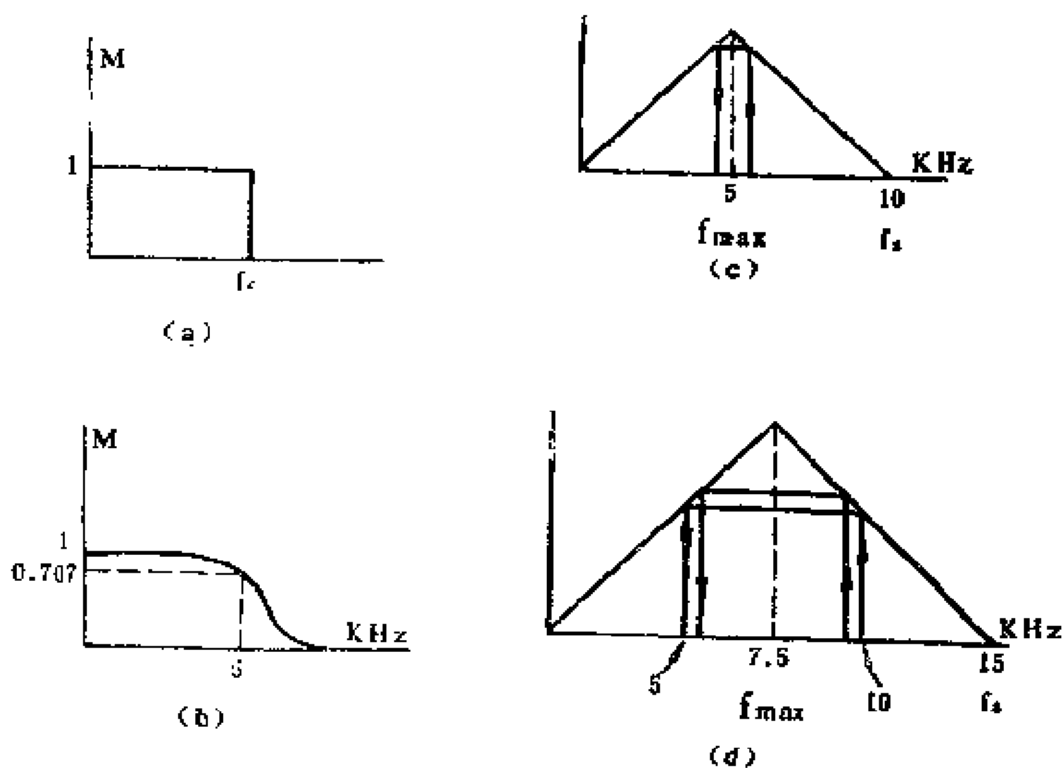


图1-3 滤波器的抗混叠作用

处，而真正被滤掉的频率成分在$f_s>f_c$以外。设图1-3b低通滤波器的标称截止频率为5kHz，用它作为抗混叠滤波器，应当取多大的Δ？采样频率取为$f_S=2\times5=10$kHz（每秒采10000个数据）

行不行？由于高于 $5kH_z$ 的频率成分在滤波后依然存在，并反映到0-5kH_z范围中来，因此应当增大采样频率，例如取$f_s=\frac{1}{\Delta}=15kH_z$（即折叠频率为7.5$kH_z$，如图1-3$d$）。这时，只有高于10$kH_z$的频率成分才会反映到0-5$kH_z$范围内，而大于10$kH_z$的成分经滤波器后已被充分衰减了，因此频率混叠效应被大大削弱。

关于采样和频率混叠的重要性质归纳如下：

（1）能根据采样信号复现原连续信号所必须的最小采样频率应当是$2f_{max}$，其中f_{max}是原来信号的频率范围，即

$$f_{s\,min}=(1/\Delta)_{min}=2f_{max} \tag{1-1-5}$$

或

$$\Delta\leqslant 1/(2f_{max}) \tag{1-1-6}$$

换言之，每周两次的采样频率可以近似得出原来的信号，这实际上就是申农采样定理。当然，如果采样点不在振荡的峰点，要重建这个信号还是有较大的误差，因而实际上以每周2.5次的采样速度作为下限。

（2）如果采样频率为$1/\Delta$，而信号的频率高于$1/(2\Delta)H_z$，那么会有较低的频率出现在频率分析中，这种现象就是"频率混叠"。

通常的做法是：先选定所感兴趣的最高频率f_{max}，在采样之前，先将信号通过低通滤波器移去高于f_{max}的频率成分，以防止混叠，然后根据采样定理选择适当的Δ。

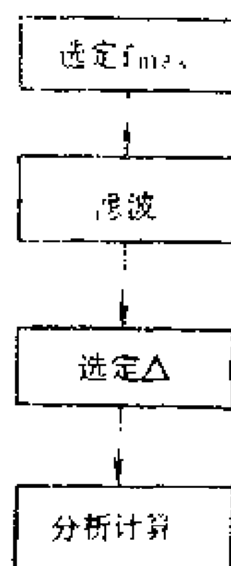


图1-4 采样间隔的选定

1.2 均值、方差和概率直方图

均值（算术均值）和方差是数据信息的两个最基本的数字特征。由于只能得到有限的随机变量的样本，而不是时间无限长的记录，因此我们要用样本均值和样本方差来对总体参数作估计。对于采样数据，样本均值为

$$m=\bar{x}=\frac{1}{N}\sum_{i=0}^{N-1}x(i) \tag{1-2-1}$$

其中$x(i)$为第i个采样值，N为采样点总数。

无偏的样本方差为

$$s^2=\frac{1}{N-1}\sum_{i=0}^{N-1}(x(i)-\bar{x})^2 \tag{1-2-2}$$

在统计学参考书中可以找到样本均值和样本方差的性质及其置信区间的讨论。对于时间序列通常所具有的记录长度而言，以上二参数的统计波动较小，因此上述估计不成问题。

如果在s^2的表达式中将$1/(N-1)$改为$1/N$，则s^2的期望值为$((N-1)/N)\sigma^2$，从而估计是有偏的（这里σ^2为总体的方差），但它是渐近无偏的，通常在时间序列分析中用到的N至少有100，甚至大到10000，所以用N来除所引起的s^2的偏差并不显著。

均值和方差的计算在原则上很简单，但在数字计算机上实现有时会存在问题。例如，由离散数据计算样本均值的直接算法公式为

$$\bar{x}_K=[\sum_{i=0}^{k-1}x(i)]+x(k) \quad k=1,\ 2,\ \cdots,\ N-1 \tag{1-2-3}$$

$$\bar{x}=\frac{1}{N}\bar{x}_{N-1}$$

亦即按顺序将数据累加，然后平均。这种方法运算量最省，但可能发生这样的问题：k个值的和可能比 $x(k)$大得多，以致舍入误差和$x(k)$的幅值为同一数量级。较为精确的算法是按下列方式计算部分和。假定数据点数是2的某次方，即$N=2^m$，则有m个计算阶段：

$$\bar{x}_1(i)=\frac{1}{2}[x(2i)+x(2i+1)] \quad i=0,\ 1,\ \cdots,\ \frac{N}{2}-1$$

$$\bar{x}_2(i)=\frac{1}{2}[\bar{x}_1(2i)+\bar{x}_1(2i+1)] \quad i=0,\ 1,\ \cdots,\frac{N}{4}-1$$

$$\vdots \tag{1-2-4}$$

$$\bar{x}=\bar{x}_m(0)=\frac{1}{2}[\bar{x}_{m-1}(0)+\bar{x}_{m-1}(1)]$$

例如当$N=8(m=3)$时，有数据$x(0),x(1)$，…，$x(7)$，共须三段计算：

（1）计算

$$\bar{x}_1(0)=\frac{1}{2}[x(0)+x(1)]$$

$$\bar{x}_1(1)=\frac{1}{2}[x(2)+x(3)]$$

$$\bar{x}_1(2)=\frac{1}{2}[x(4)+x(5)]$$

$$\bar{x}_1(3)=\frac{1}{2}[x(6)+x(7)]$$

（2）计算

$$\bar{x}_2(0)=\frac{1}{2}[\bar{x}_1(0)+\bar{x}_1(1)]$$

$$\bar{x}_2(1)=\frac{1}{2}[\bar{x}_1(2)+\bar{x}_1(3)]$$

(3)最后计算样本均值

$$\bar{x}=\bar{x}_3(0)=\frac{1}{2}[\bar{x}_2(0)+\bar{x}_2(1)]$$

这种多级均值算法可以克服舍入误差的影响。类似方法可以用于任何序列求和，当然同样可推广用来计算平方和或互积和。

方差的计算也有不同的算法，例如有移去均值算法

$$s_x{}^2=\frac{1}{N-1}\sum_{i=0}^{N-1}[x(i)-\bar{x}]^2 \qquad (1\text{-}2\text{-}5)$$

和直接算法

$$s_x^2=\frac{1}{N-1}[(\sum_{i=0}^{N-1}x^2(i))-N\bar{x}^2] \qquad (1\text{-}2\text{-}6)$$

直接算法的问题是其中$\Sigma x^2(i)$和$N\bar{x}^2$可能都是很大的数，而它们之差却很小，这时式(1-2-6)右边差值的舍入误差可能和s_x^2具有同一数量级。移去均值算法对舍入误差较不敏感，每个数据减去$\bar{x}$，使差值的"动态范围"减小，这使求和的结果受舍入误差的影响降低了。

下面我们将介绍样本概率密度或直方图，它们可以帮助分析数据的分布性质，同时也有助于判断数据是否"正常"或"合理"。

和$\bar{x}$，s^2不同，对于一组给定的数据，样本密度函数并不是唯一的，它还取决于求取时所用的某些参数。直方图是这样计算的，将x的一个区间范围（例如$a<x<b$）分成k个相等的子区间（或组区间），这样整个x的范围看成由$(k+2)$个区间构成（图1-5)。考察全部数据并记录其在每一区间内出现的次数。以$\{N_j\}$表示在第j个区间出现$\{x(i)\}$的次数的集合，令$c=(b-a)/k$，$d_j=a+jc$，则$\{N_j\}$可列表如下

表 1-1

j	N_j
0	$x<a$出现的次数
·	·
·	·
i	$d_{j-1}\leqslant x<d_j$出现的次数
·	·
·	·
k	$d_{k-1}\leqslant x<b$出现的次数
k+1	$x\geqslant b$出现的次数

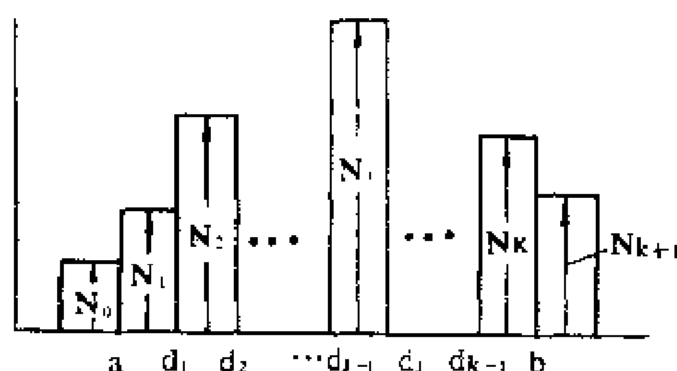


图1-5 概率直方图

在计算机上进行上述分类的具体做法是按以下规则依次检查每个$x_i(i=1,\ 2,\ \cdots,\ N)$

(1)若$x(i)<a$则使N_0加1。

(2)若$x(i)\geqslant b$则使N_{k+1}加1。

(3)若上述二条件均不满足，则$a\leqslant x(i)<b$计算

$$j=\lfloor\frac{x(i)-a}{c}\rfloor+1 \qquad (1-2-7)$$

我们用 $I=\lfloor A\rfloor$表示：对于任一实数A，整数I是实数 A 的整数部分，即I是小于或等于A的最大整数。由式(1-2-7)求得j后使N_j加1。

以上方法在计算机上是很容易实现的。利用上述结果可以画出直方图，即 $\{N_j\}$ 序列，其次是 $\{P_j\}$ 序列。这里

$$P_j = 使 (d_{j-1} \leqslant x < d_j) 的样本概率$$
$$= N_j / N \qquad j = 0, 1, 2, \cdots, k+1 \qquad (1\text{-}2\text{-}8)$$

此外，还可以算出样本概率密度函数(PDF)，即 $\{p_j\}$ 序列 ($j=0, 1, \cdots, k+1$)

$$p_j = \frac{P_j}{c} = \frac{N_j/N}{c} = \frac{N_j k}{N(b-a)} \qquad (1\text{-}2\text{-}9)$$

它可以看作是在每一区间中点处分布函数的导数。

在采取上述步骤之前先要选择a, b和k，那么它们的选择原则是什么？这没有唯一的答案，主要是根据对所考察数据分布规律的假定以及收集数据的方法。在上面讨论中，实际上相当于假定数据是局限在一定范围内，而且在该范围内只有有限个可能的数据级别。例如，当模数转换器只有128个量级，则取$k>128$显然没有意义，因为有些量级必然是空的，徒然增加不必要的计算量。

另一方面，也可以看到PDF子区的划分也会出现不正确的结果。例如采用16个量级的模数转换器，各量级(0, 1, …15)均为等间隔，因而第i个计数出现的先验概率为1/16。如果$k=12$，$a=0$，$b=15$，则出现的分布在表1-2中列出。可见，在1，5，9三组中包含的数据出现的可能性是其它组的两倍，所以在这些地方PDF值将有较大的偏差。

上面例子告诉我们，当计算PDF时要注意根据数据的来源，对其总体作合理的假设进行适当的分组，关于确定a，b，k诸参数的一种准则将在下一节关于正态性检验问题中提出。

检查数据是否“合理”或“正常”的方法，除了查看全部记录曲线（这是很不上算的），经济有效的办法是求它的样本PDF，通过分析可以得到有关数据性质的许多信息，并发现可能存在的问题。

表1·2

K	范　围	包含的数据量级
1	0-1.25	0,1
2	1.25-2.50	2
3	2.50-3.75	3
4	3.75-5.00	4
5	5.00-6.25	5,6
6	6.25-7.50	7
7	7.50 8.75	8
8	8.75-10.00	9
9	10.00-11.25	10,11
10	11.25-12.50	12
11	12.50-13.75	13
12	13.75-15.00	14

图1-6表示有两个异点的数据波形，相应的样本PDF示于图1-7。可以看出，大量的数据都密集在一起，而两个异点，则远离右边，中间的区组又都是空的，这种具有异点的数据由于数据范围很大，迫使正常数据都集中到一两个区组中去。

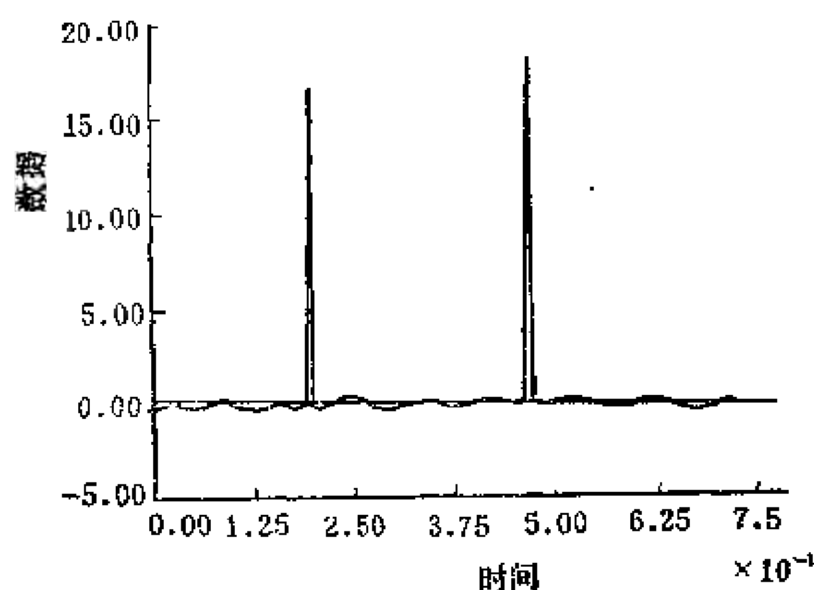


图1-6　含异点的序列

```
   WILD POINT PDF
NDATA =    201
 XBEC -   -3.000   XEND =   65.000   DELTAX =    2.000
 MEAN =     .621   STANDARD DEVIATION =    6.586

 -4.000   1.0        *
 -2.000  33.0   xxxxx*xx
   .000131.0    xxxxx*xxxxxxxxxxxxxxxxxxxxxxxxx
  2.000  33.0   xxxxx*xx
  4.000   1.0        *
  6.000    .0       *
  8.000    .0      *
 10.000    .0     *
 12.000    .0    *
 14.000    .0    *
 16.000    .0
 18.000    .0
 20.000    .0
 22.000    .0
 24.000    .0
 26.000    .0
 28.000    .0
 30.000    .0
 32.000    .0
 34.000    .0
 36.000    .0
 38.000    .0
 40.000    .0
 42.000    .0
 44.000    .0
 46.000    .0
 48.000    .0
 50.000    .0
 52.000    .0
 54.000    .0
 56.000    .0
 58.000    .0
 60.000   1.0
 62.000    .0
 64.000    .0
 66.000   1.0
```

图1-7 有异点的概率直方图

图1-8是另一种数据的样本PDF，数据在两端区出现的次数多，而在两端之外又没有数据，这是“箝位”的征兆。产生这种情况有不同的原因，最常见的是模拟放大器运行在其线性范围之

外，或是所设置的模数转换器没能完全复盖整个数据的变化范围。

```
U
    DATA OF CW
NDATA =    100
 XBEG =    -.700    XEND =     .900  DELTAX =     .100
 MEAN =     .051    STANDARD DEVIATION =     .571


  -.750    .0     *
  -.650  27.0   xx*xxxxxxxxxxxxxxxxxxxxxxxx
  -.550   4.0   xxx*
  -.450   1.0   x    *
  -.350   2.0   xx   *
  -.250   2.0   xx     *
  -.150   6.0   xxxxxx*
  -.050   2.0   xx       *
   .050   5.0   xxxxx *
   .150   3.0   xxx     *
   .250   7.0   xxxxxx*
   .350   4.0   xxxx *
   .450   2.0   xx   *
   .550   2.0   xx   *
   .650   5.0   xxx*x
   .750  28.0   xx*xxxxxxxxxxxxxxxxxxxxxxxxx
   .850    .0     *
   .950     0    *
 CHI SQUARE = 381.857
```

图1-8 被箝位的序列的概率直方图

关于分级计算样本均值，移去均值计算样本方差，概率直方图的计算、绘图，以及样本概率密度的计算，在本书附有源程序SMSV和PDFH可供使用。

1.3 随机数据的正态性检验

正态分布的PDF（记作p）为

$$p(x)=(2\pi\sigma^2)^{-\frac{1}{2}}\exp(-(x-\mu)^2/(2\sigma^2)) \qquad (1-3-1)$$

式(1-3-1)中，μ和σ^2分别为总体的均值和方差。概率分布函数是密度函数的积分

$$P(x<X)=(2\pi\sigma^2)^{-\frac{1}{2}}\int_{-\infty}^{X}\exp(-(t-\mu)^2/2\sigma^2)dt$$

$$=(2\pi)^{-\frac{1}{2}}\int_{-\infty}^{(x-\mu)/\sigma}\exp(-\frac{1}{2}t^2)dt$$

$$=\Phi((x-\mu)/\sigma) \tag{1-3-2}$$

其中Φ称为“概率积分”。随机变量处于α和β之间的概率为

$$P(\alpha\leqslant x<\beta)=\Phi((\beta-\mu)/\sigma)-\Phi((\alpha-\mu)/\sigma) \tag{1-3-3}$$

在时序分析中经常有正态分布的假定，有些理论也建立在正态假定的基础上，所以常常需要知道采集的数据究竟是否是正态的。

“χ^2 拟合优度检验”是一种可取的方法，它是利用 χ^2 统计量作为观察到的PDF和理论密度函数之间偏差的量度，两者是否相同可通过分析χ^2的样本分布来检验。如果数据是正态的，则应落入第j组区间中的数据个数（称为组区间中的期望频数）为

$$F_0=N\Phi(\frac{a-m}{s})$$

$$\vdots$$

$$F_j=N[\Phi(\frac{a+jc-m}{s})-\Phi(\frac{a+c(j-1)-m}{s})]$$

$$\vdots \tag{1-3-4}$$

$$F_{k+1}=N[1-\Phi(\frac{b-m}{s})]$$

F_j和观察到的频数偏差为N_j-F_j，显然

$$\sum_{j=0}^{k+1}N_j=\sum_{j=0}^{k+1}F_j=N \tag{1-3-5}$$

故总的偏差必为0，根据皮尔逊定理[1]，样本的χ^2变量为

$$\chi^2=\sum_{j=0}^{k+1}(N_j-F_j)^2/F_j \tag{1-3-6}$$

假定这个样本χ^2变量近似为χ^2分布，可以将它和理论的χ^2分布(记作$\chi^2_{n;\alpha}$)作比较，这时自由度n等于$(k+2)$(如果把范围两端的组也算上的话)减去一些线性约束的数目——一个约束是：当前$(k+1)$个组区间的频数已知时，由于总频数为N，最后一个组区间的频数也就知道了；另外两个约束是由于同理论正态密度函数拟合观察数据的频数直方图而引起的，这就是用样本均值和样本方差而不是真正的均值和方差来计算$\{F_i\}$。因此，如果利用全部$\{N_i\}$，则自由度

$$n=(k+2)-3=k-1 \tag{1-3-7}$$

实际n值可能比这还要小些，因为$F<2$的一些组可能和其它组合并。

在X^2的自由度正确确定之后，可以作如下假设检验：若假设x变量是正态的，在把观察数据分组列入$k+2$个组区间后，利用样本均值和方差计算F_i，求出的X^2样本PDF对正态分布的任何偏离都会使X^2增大。如果

$$X^2\leqslant\chi^2_{n;\alpha} \tag{1-3-8}$$

则接受(在α显著水平上)数据为正态分布的假设；反之，如果X^2大于$\chi^2_{n;\alpha}$，则在α显著水平上拒绝上述假设。常用的显著水平为5%，10%和20%，即相应于95%，90%和80%的置信水平。

对于用χ^2拟合优度检验检查正态性的组区间数目，肯德尔(*Kendall*)和斯图尔特(*Stuart*)给出了正态总体样本量和分组数目应满足的最优关系式：

$$\text{分组区间数目}=1.87(N-1)^{2/5} \tag{1-3-9}$$

该式中假定数据是无关的，且$\alpha=0.05$，结果列于表1-3。

在应用χ^2检验时的一个准则是每个区间中的期望频数至少应为2。由于范围两端的期望频数最少，因此上述要求可以用来确定a和b。参数a应满足(由式(1-3-2)中$P=2/N$)：

$$2=N\left[(2\pi)^{-\frac{1}{2}}\int_{-\infty}^{(a-m)/s}\exp\left(-\frac{1}{2}t^{2}\right)dt\right] \quad (1\text{-}3\text{-}10)$$

由该隐式可以求得a，又由平均值$m=(b-a)/2$得参数b为

$$b=2m-a \quad (1\text{-}3\text{-}11)$$

参数k为

$$k=(\text{分组区间数目})-2=r-2 \quad (1\text{-}3\text{-}12)$$

这三个参数确定之后就可以计算样本PDF。

表1-3　样本量为N，$\alpha=0.05$时最小的分组区间数目(r)

N	r	N	r
200	16	20,000	94
400	20	40,000	129
600	24	70,000	162
800	27	100,000	187
1,000	30	200,000	247
1,500	35	400,000	326
2,000	39	700,000	407
4,000	57	1,000,000	470
7,000	65	1,140,000	500
10,000	74		

例：图1-9是由计算机产生的均匀分布白噪声（用DAGENT程序）1001个采样点得出的样本PDF。理论上PDF为

$$f(x)=\begin{cases}1 & -\frac{1}{2}\leqslant x<\frac{1}{2}\\ 0 & \text{其它}\end{cases} \quad (1\text{-}3\text{-}13)$$

故理论上的均值和方差为

$$\mu=\int_{-\infty}^{\infty}xf(x)dx=0 \quad (1\text{-}3\text{-}14)$$

```
   UNIFORM PDF
NDATA =   1001
 XBEG =   -1.000    XEND =    1.000   DELTAX =     .050
 MEAN =     .021    STANDARD DEVIATION =     .576

 -1.025    .0        *
  -.975  21.0   xxxxxxx*xxxxxxxxxxxxx
  -.925  26.0   xxxxxxxx*xxxxxxxxxxxxxxxxx
  -.875  22.0   xxxxxxxxx*xxxxxxxxxxxx
  -.825  29.0   xxxxxxxxxxx*xxxxxxxxxxxxxxxxx
  -.775  32.0   xxxxxxxxxxxx*xxxxxxxxxxxxxxxxxxx
  -.725  18.0   xxxxxxxxxxxxxx*xxx
  -.675  19.0   xxxxxxxxxxxxxxxx*xx
  -.625  21.0   xxxxxxxxxxxxxxxxx*xxx
  -.575  25.0   xxxxxxxxxxxxxxxxxxx*xxxxx
  -.525  14.0   xxxxxxxxxxxxxx       *
  -.475  27.0   xxxxxxxxxxxxxxxxxxxxxxx*xxx
  -.425  29.0   xxxxxxxxxxxxxxxxxxxxxxxxx*xxx
  -.375  22.0   xxxxxxxxxxxxxxxxxxxxxx    *
  -.325  23.0   xxxxxxxxxxxxxxxxxxxxxxx     *
  -.275  23.0   xxxxxxxxxxxxxxxxxxxxxxx      *
  -.225  29.0   xxxxxxxxxxxxxxxxxxxxxxxxxxxxx  *
  -.175  25.0   xxxxxxxxxxxxxxxxxxxxxxxxx       *
  -.125  25.0   xxxxxxxxxxxxxxxxxxxxxxxxx        *
  -.075  36.0   xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx*xx
  -.025  31.0   xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx   *
   .025  20.0   xxxxxxxxxxxxxxxxxxxx              *
   .075  21.0   xxxxxxxxxxxxxxxxxxxxx             *
   .125  26.0   xxxxxxxxxxxxxxxxxxxxxxxxxx       *
   .175  18.0   xxxxxxxxxxxxxxxxxx              *
   .225  22.0   xxxxxxxxxxxxxxxxxxxxxx          *
   .275  38.0   xxxxxxxxxxxxxxxxxxxxxxxxxxxxxx*xxxxxxx
   .325  30.0   xxxxxxxxxxxxxxxxxxxxxxxxxxxxx*
   .375  23.0   xxxxxxxxxxxxxxxxxxxxxxx     *
   .425  20.0   xxxxxxxxxxxxxxxxxxxx      *
   .475  23.0   xxxxxxxxxxxxxxxxxxxxxxx *
   .525  23.0   xxxxxxxxxxxxxxxxxxxxxxx*
   .575  25.0   xxxxxxxxxxxxxxxxxxxxx*xxx
   .625  28.0   xxxxxxxxxxxxxxxxxxx*xxxxxxxx
   .675  23.0   xxxxxxxxxxxxxxxxx*xxxxx
   .725  34.0   xxxxxxxxxxxxxxx*xxxxxxxxxxxxxxxxxx
   .775  24.0   xxxxxxxxxxxxxx*xxxxxxxxx
   .825  27.0   xxxxxxxxxxxx*xxxxxxxxxxxxxx
   .875  32.0   xxxxxxxxxxx*xxxxxxxxxxxxxxxxxxxx
   .925  26.0   xxxxxxxxx*xxxxxxxxxxxxxxxx
   .975  21.0   xxxxxxxx*xxxxxxxxxxxx
  1.025    .0          *
 CHI SQUARE = 299.641
```

图1-9　均匀分布序列的概率直方图

$$\sigma^2=\int_{-\infty}^{\infty}(x-\mu)^2 f(x)dx=\int_{-\frac{1}{2}}^{\frac{1}{2}}x^2dx=\frac{x^3}{3}\Big|_{-\frac{1}{2}}^{\frac{1}{2}}$$

$$=1/12=0.083333 \tag{1-3-15}$$

样本PDF的值与此很接近。如果按样本的均值和方差作正态分布曲线，显示二者不是很接近的。事实上，用χ^2检验也证明拟合失败。

图1-10也是伪随机数产生的噪声，其理论的均值和方差分别为0，1，得到的样本PDF的值如图上所注。该数据以95％的置信水平可以认为是正态分布。

```
     NORMAL PDF
 NDATA  =    1001
 XBEG  =   -3.000    XEND =    3.000   DELTAX =     .150
 MEAN  =    -.032    STANDARD DEVIATION =   1.015

 -3.075   1.0   x
 -2.925    .0   *
 -2.775    .0   *
 -2.625   5.0   *xx
 -2.475   6.0   x*x
 -2.325   5.0   x*x
 -2.175   6.0   xx*
 -2.025   8.0   xxx*
 -1.875  13.0   xxxxx*x
 -1.725  10.0   xxxxx *
 -1.575  19.0   xxxxxxxx*x
 -1.425  19.0   xxxxxxxxxx*
 -1.275  31.0   xxxxxxxxxxxxx*xx
 -1.125  30.0   xxxxxxxxxxxxxxx  *
  -.975  37.0   xxxxxxxxxxxxxxxxxxx*
  -.825  45.0   xxxxxxxxxxxxxxxxxxxxxx*x
  -.675  53.0   xxxxxxxxxxxxxxxxxxxxxxxx*xxx
  -.525  55.0   xxxxxxxxxxxxxxxxxxxxxxxxxx*xx
  -.375  46.0   xxxxxxxxxxxxxxxxxxxxxxxx    *
  -.225  67.0   xxxxxxxxxxxxxxxxxxxxxxxxxxxxx*xxxxx
  -.075  66.0   xxxxxxxxxxxxxxxxxxxxxxxxxxxxx*xxxx
   .075  54.0   xxxxxxxxxxxxxxxxxxxxxxxxxxxx *
   .225  58.0   xxxxxxxxxxxxxxxxxxxxxxxxxxxxx
   .375  57.0   xxxxxxxxxxxxxxxxxxxxxxxxxxx*xx
   .525  48.0   xxxxxxxxxxxxxxxxxxxxxxxxx*
   .675  46.0   xxxxxxxxxxxxxxxxxxxxxxx*
   .825  37.0   xxxxxxxxxxxxxxxxxxxx *
    975  36.0   xxxxxxxxxxxxxxxxxx*
  1.125  39.0   xxxxxxxxxxxxxxx*xxxxx
  1.275  30.0   xxxxxxxxxxxxx*xx
  1.425  16.0   xxxxxxxx  *
  1.575  14.0   xxxxxxx*
  1.725   9.0   xxxxx *
  1.875   7.0   xxxx*
  2.025   7.0   xxx*
  2.175   5 0   xx*
  2.325   1.0   x*
  2.475   5.0   *xx
  2.625   2.0   *
  2.775   1.0   *
  2.925   3.0   xx
  3.075   4.0   xx
 CHI SQUARE =   55 771
```

图1-10 正态分布的概率直方图

1.4 随机数据的独立性检验

在时序分析和建模过程中，许多场合都要对某一随机过程是否独立进行检验。现有的方法和准则也有好几种，本节给出的两种方法是基于正态随机变量自相关函数的统计性质。

对于正态独立分布的随机变量 $x \sim NID(0, \sigma_x^2)$ ，其自相关函数

$$\rho(r)=\delta(r)=\begin{cases}1 & r=0\\ 0 & r\neq 0\end{cases} \tag{1-4-1}$$

因此，当$r\geqslant 1$时$\rho(r)=0$。但我们所能得到的是样本自相关$\hat{\rho}(r)$，它一般不会等于$\delta(r)$。如何从估计的相关函数判断“真正”的函数是否满足独立性条件，可以利用巴特利公式（见第三章3.1.2节），该公式指出：若$\rho(r)$ 在 $r>M$时趋于零，则当N足够大的情况下

$$Var[\hat{\rho}(r)]\simeq\frac{1}{N}\sum_{m=-M}^{M}\rho^2(m) \quad r>M \tag{1-4-2}$$

而且$\hat{\rho}(r)$ （$r>M$）近似正态分布。

显然若$\hat{\rho}(r)$是白噪声的相关函数，则$M=0$

$$Var[\hat{\rho}(r)]\simeq 1/N \quad r>0 \tag{1-4-3}$$

而若

$$|\hat{\rho}(r)|\leqslant 1.96\sqrt{1/N}\simeq 2\sqrt{1/N} \tag{1-4-4}$$

或

$$\sqrt{N}\,|\hat{\rho}(r)|\leqslant 2 \tag{1-4-5}$$

便可认为$\hat{\rho}(r)$为零的可能性是95%，从而接受$\rho(r)=0(r>0)$这一估计，即数据是独立的。

检验x是否独立的另一种方法是波克斯(*Box*)和皮尔斯（*Pie-*

rce)提出的“一揽子”或整体检验。考虑到$r \geqslant 1$时白噪声序列的样本自相关的分布渐近于正态，即

$$\hat{\rho}(r) \sim \mathrm{NID}(0,\ 1/N)$$

故$\hat{\rho}(r)/\sqrt{1/N}$，或是说$\{\sqrt{N}\hat{\rho}(1),\ \sqrt{N}\hat{\rho}(2), \cdots, \sqrt{N}\hat{\rho}(k)\}$当$N$较大时，这$k$个量近似为相互独立的正态NID(0,1)的随机变量，因而它们的平方和符合χ^2分布。若构造统计量

$$Q = N\sum_{r=1}^{k}\hat{\rho}^2(r) \tag{1-4-6}$$

则检验x_1，x_2，…，x_N是否为白噪声的样本值的问题，转化为检验统计量Q是否是自由度为k的χ^2分布的问题。

可取$k=20$，以“$\{x_t\}$为白噪声”做原假设，以α为显著性水平，根据α和自由度k由χ^2－分布表查出相应的$\chi^2_\alpha(k)$值，并与计算出的Q值比较。当

$$Q \leqslant \chi^2_\alpha(k) \tag{1-4-7}$$

时则肯定原假设，即在$(100-\alpha)\%$的置信水平上接受$\{x_t\}$为独立的假定。若

$$Q > \chi^2_\alpha(k) \tag{1-4-8}$$

则否定原假设。

表1-4　$(100-\alpha)\%$水平上的$\chi^2_\alpha(20)$

α	10	5	1	0.1
临界点	28.41	31.41	37.57	45.31

独立性检验的程序见附录二的TESTI。

1.5　非平稳趋势的检验

时间序列分析的一些主要方法都是假定数据样本是来自平稳和各态历经的随机过程，也就是它们的期望（均值、方差、相关

等）都不随时间推移而变化，而且可以用时间平均代替总体平均。当然，这样做的时候应当有足够长的数据记录才能使样本平均有代表意义。当任何一种平稳性条件被破坏时就出现非平稳。在实际情况中经常出现的有三种非平稳过程——均值非平稳，方差非平稳，以及均值方差非平稳。

维纳（Wiener，1930）曾经分析过的一种非平稳过程（即称维纳过程），可以看成是一般的平稳过程经过积分器（如在$t=0$开始）所得的结果，以x为输入，y为输出

$$y(t)=\int_0^t x(\tau)d\tau \tag{1-5-1}$$

若$E(x)=\mu_x=\text{const}$，则$E(y)$随时间变化

$$\mu_y(t)=\begin{cases} t\mu_x & t\geqslant 0 \\ 0 & t<0 \end{cases} \tag{1-5-2}$$

若x的相关函数为$R_x(\tau)$，则可证明y的方差为

$$\sigma_y^2(t)=\begin{cases} \int_0^t\int_0^t R_x(t_1-t_2)dt_1dt_2 & t\geqslant 0 \\ 0 & t<0 \end{cases} \tag{1-5-3}$$

当$R_x(\tau)=\sigma_x^2\delta(\tau)$（白噪声时）

$$\sigma_y^2(t)=t\sigma_x^2 \tag{1-5-4}$$

实际中，根据加速度传感器的数据经积分求速度就是这种例子。如果加速度项带有噪声，则积分结果的噪声就有上面类似的趋势出现。

这里介绍一种对均值或方差可能存在某种趋势进行检验的方法。首先由时间序列求出一个大致不相关的均值或方差值的序列（这可以把整个数据记录分成M段，然后求各段按时间平均的均值和方差）。设该序列为y_1，y_2，…，y_M。每当出现$y_j>y_i$（$j>i$，$i=1$，2，…，$M-1$）时定义为y_i的一个逆序，对于下标为i的已

知值y_i，其逆序数定义为与y_i相应的逆序的个数A_i。逆序总数为

$$A=\sum_{i=1}^{M-1} A_i \tag{1-5-5}$$

例如序列2，3，2，4，5，3，对于$y_1=2$有$A_1=4$（因y_2，y_4，y_5，y_6均大于y_1，对于$y_2=3$有$A_2=2$（因y_4，y_5均大于y_2），还可看出$A_3=3$，$A_4=1$，$A_5=0$，总数$A=10$。可以证明，以随机整数序列出现的A的平均值为

$$E(A)=M(M-1)/4 \tag{1-5-6}$$

式（1-5-6）的证明如下：

y_1比随后的$M-1$个随机数大或小的可能性相同，故由这种比较而得的逆序数的平均值为$E(A_1)=(M-1)/2$，将y_2与余下的$M-2$个数比较有$E(A_2)=(M-2)/2$，同理有$E(A_{M-2})=2/2$，$E(A_{M-1})=1/2$。于是

$$E(A)=\sum_{i=1}^{M-1} E(A_i)=\frac{1}{2}\sum_{i=1}^{M-1} i=M(M-1)/4 \tag{1-5-7}$$

还可以证明

$$Var[A]=M(2M^2+3M-5)/72 \tag{1-5-8}$$

为节省篇幅，式(1-5-8)的证明从略。统计量

$$u=(A+\frac{1}{2}-E(A))/\sqrt{Var(A)} \tag{1-5-9}$$

渐近服从正态分布$N(0, 1)$。根据M算出$E(A)$，然后按实际逆序数A得出u值。如果u值是处在± 2之内，则可接受“序列无趋势”的假设；否则拒绝该假设（在0.05显著水平上）。显然，如A很大，表明序列均值（或方差）有上升的趋势；而A很小，则表明序列的均值（或方差）有下降的趋势。

上述非平稳趋势检验对于单调的趋势是有效的，但在有些情况下具有局限性。例如序列方差非平稳形如图1-11所示，在正跳



图1-11 方差变化图之一

部分有很多逆序，但在随后的负跳部分则逆序很少，所以整个数据的逆序检验却可以是正常的。类似的道理，对图1-12*a*，*b*，*c*形式的方差非平稳性检验可能有效，而对于图1-12*d*，*e*，*f* 形式的非平稳性检验则可能无效。

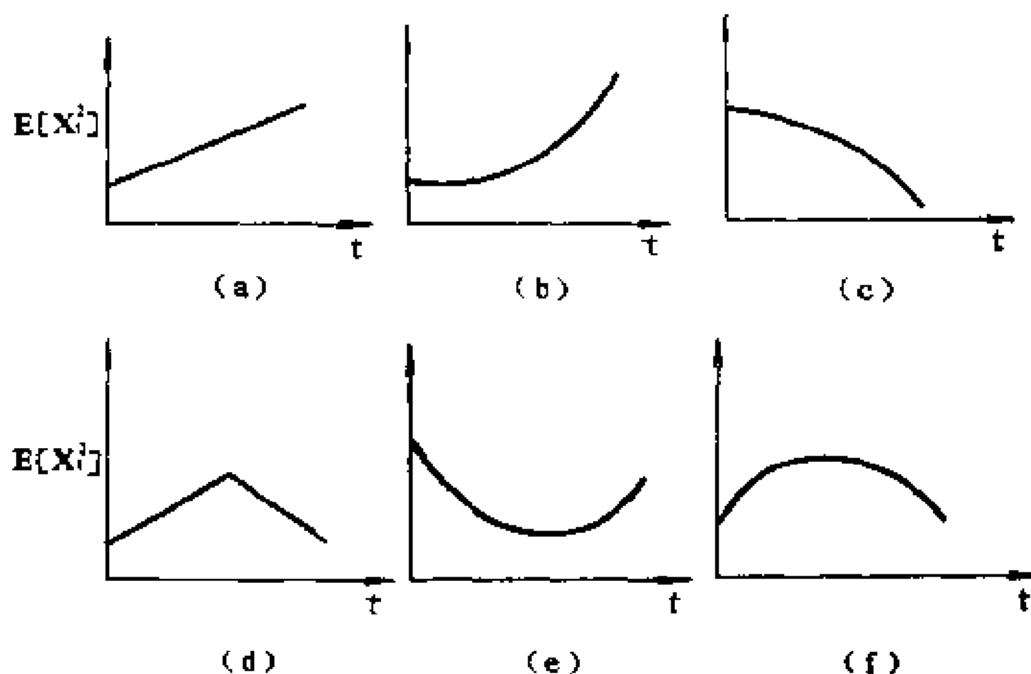


图1-12 方差变化图之二

检验序列中是否有潜在趋势的一种非参数方法是“游程检验”（轮次检验）。在保持序列原有顺序的情况下，游程定义为具有相同记号的序列，这种记号把观察值分成两个互相排斥的类。例如观察序列的值是 $x_i(i=1, 2, \cdots, N)$，其均值为 $\bar{x}$，用

记号“+”表示$x_i \geqslant \bar{x}$，“-”表示$x_i < \bar{x}$。按符号“+”和“-”的出现顺序将原序列写成如下形式。例如

$$\underbrace{+\ +\ +}_{1}\ \underbrace{-}_{2}\ \underbrace{+\ +}_{3}\ \underbrace{-\ -}_{4}\ \underbrace{+}_{5}\ \underbrace{-\ -\ -\ -}_{6}\ \underbrace{+}_{7}$$

共有“+”号和“-”号14个，分7个游程。每个游程的长短在这里并不重要。游程太多或太少都被认为是存在非随机性趋势。游程检验所判断的原假设为：“样本数据出现的顺序没有明显的趋势”。我们采用的样本统计量有

N_1=一种记号出现的总数。

N_2=另一种记号出现的总数。

r=游程的总数。

其中r作为检验统计量，把N_1和N_2小于或等于15认为是小样本量；否则为大样本量。对于显著水平$\alpha=0.05$的双边检验，表附Ⅲ-4给出概率分布左右两侧为$\alpha/2=0.025$时的上限r_U和下限r_L。如果r在界限以内则接受原假设，否则拒绝原假设。

例如有$N=22$的观察序列，其值超过均值者记为“+”；反之记为“-”，得符号序列如下

$$++---+--+++++--+--++-+$$

我们按$\alpha=0.05$检验顺序的随机性。因$N_1=12(+)$，$N_2=10(-)$，$r=11$。查表附Ⅲ-4得原假设的接受域为$7 \leqslant r \leqslant 17$，故原序列没有明显的潜在趋势。

当N_1或N_2超过15时可认为是大样本量，这时可以用正态来近似，即可利用正态分布表（表附Ⅲ-1）来定出检验的接受域和否定域。统计量为

$$z=\frac{\text{游程数}-\text{游程的期望数}}{\text{游程标准差}}=\frac{r-\mu_r}{\sigma_r} \qquad (1\text{-}5\text{-}10)$$

其中

$$\mu_r = \frac{2N_1N_2}{N} + 1$$

$$\sigma_r = \left[\frac{2N_1N_2(2N_1N_2 - N)}{N^2(N-1)}\right]^{1/2}$$

$$N = N_1 + N_2$$

对于$\alpha=0.05$的显著水平，如果$|z|\leqslant 1.96$，则可接受原假设。

1.6 剔点处理

数据采集系统在工作中有时会引入一些虚假数据，其原因可能是传输线中信号的丢失、数字化装置的故障或差错等所造成。由此而产生的异点，在随后的分析中会造成许多问题。例如一个达到数字化装置最大值的异点会使谱分析的整个噪声水平增大，而两个相距很近的这种异点将在谱分析中产生许多虚假的频率。

因此，在整个数据分析过程中，最好先进行异点的检测和剔除。但要做到恰如其份并不容易，还没有一个自动剔点的普遍方法。必要时可以设置便于人为校正的某种数字程序，数据先通过它检测出异点，并用曲线或图表形式打印出这些异点的有关信息。分析人员在查看了这些打印信息后，对下一程序设定输入信息以代替这些点或干脆排除不用。

已有的许多剔除异点的自动数字程序还不十分完善，下面列举两种。

第一种方法见图1-13。该方案采用两个数字低通滤波器，它的输出是对输入函数的平滑估计。这里隐含认为正常的数据是“平滑”的，而异点是“突变”的。其中$(\bar{x}(i))^2$是先对数据作平滑再平方得到的值。$\overline{x^2(i)}$是先对数据取平方再作平滑而得到的值。图中虚线框不断产生样本方差的更新值

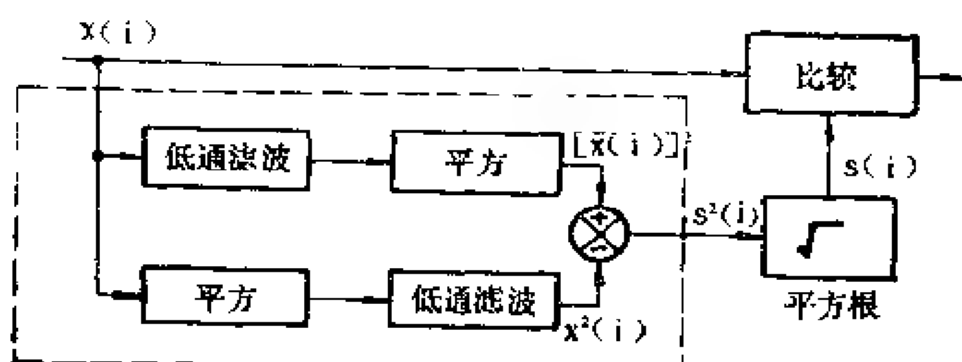


图1-13 剔点处理方案

$$s^2(i)=\overline{x^2(i)}-(\bar{x}(i))^2 \qquad (1\text{-}6\text{-}1)$$

取$s^2(i)$开方可得标准差。

接着是检查下一个数据点$x(i+1)$，如果

$$\bar{x}(i)-ks(i)<x(i+1)<\bar{x}(i)+ks(i) \qquad (1\text{-}6\text{-}2)$$

则认为$x(i+1)$是可接受的。k是分析者根据情况设定的适当数值，通常为3到9，开始时不妨取为6。

如果$x(i+1)$被认为是异点，则可用$\hat{x}(i+1)$来替代，即

$$\hat{x}(i+1)=2x(i)-x(i-1) \qquad (1\text{-}6\text{-}3)$$

这实际上是线性外推。

这种方法必须附加一些计算，即事先规定连续外推的次数，以免出现无休止的外推。因为接连检测到一些异点后，最终的外推结果可能偏离很远，以致会排除本来是正常的数据点。

第二个异点剔除方案是杜奇（*Tukey*）提出的，其基本思想是产生一个曲线的平滑估计，然后把它从数据中减掉，这样识别异点就容易得多了。该方案利用"中位数"是均值的鲁棒（*robust*）估计这个事实。其步骤为：

（1）从$x(i)$构造一个新序列$x'(i)$，方法是取$x(1)$，…，$x(5)$的中位数作为$x'(3)$，然后舍去$x(1)$加入$x(6)$取中位数得$x'(4)$，

依此类推，直到加入最后一个数据。换言之，总是在相邻的五个数据中择取中位数。显然，$x'(i)$的项数比$x(i)$项数少四项。

$$\underbrace{x(1),x(2),x(3),x(4),x(5)}_{x'(3)},x(6),\cdots,x(N-1),x(N)$$

$$\underbrace{x(2),\cdots,x(6)}_{x'(4)}\quad\cdots\quad\underbrace{x(N-4),\cdots,x(N)}_{x'(N-2)}$$

（2）用类似的方法在$x'(i)$的相邻的三个数据中择取中位数构成序列$x''(i)$

（3）最后是由序列$x''(i)$按如下方式构成序列$x'''(i)$：

$$x'''(i)=\frac{1}{4}x''(i-1)+\frac{1}{2}x''(i)+\frac{1}{4}x''(i+1) \quad (1\text{-}6\text{-}4)$$

（这是一个海宁平滑滤波器）。

分析序列$x(i)-x'''(i)$，看是否有$|x(i)-x'''(i)|>k$（预定值），如果有，则用一内插值代替$x(i)$。

1.7 提取趋势项

有时需要在某一时间序列中去掉一个线性的或缓慢变化的趋势，这种趋势项可能是由于数据中的有些分量是经过积分产生的。积分可以导致两种误差，首先是，如果零点没有调准，则在每一采样时刻都有一小误差项，经过积分后，这一常数项变成了直线。这一线性趋势在谱分析或其它计算中会导致很大的误差。

另一类误差的产生是由于积分对低频噪声起功率放大作用，

而在数据中常有这类噪声，经过积分后，变成缓慢变化的随机信号，其变化速度在某种程度上取决于采样间隔。

趋势项也并非都是误差，它可能代表时间序列中包含的有用信息，由于它的出现使过程成为非平稳。因此在对数据作平稳化预处理时也需要提取出趋势项。在本书第八章8.1，8.2两节中，我们将对如何提取趋势项，作深入讨论。

本节讨论如何提取多项式形式的趋势项。一般地说，变化着的趋势项可以用高通滤波器来消除，而多项式形式的趋势可以用最小二乘法来提取，下面我们着手讨论这种方法。

设$x(i)$是等间隔Δ采样的函数，即

$$x(i) \quad i=0,\ 1,\ \cdots,\ N-1$$

要用多项式$\hat{x}(i)$来拟合它，令

$$\hat{x}(i)=\sum_{k=0}^{K}(i\Delta)^{k}c_{k} \tag{1-7-1}$$

$\hat{x}(i)$点的集合应当是$x(i)$中多项式成分的估计。定义多项式系数的函数$E(c)$为

$$E(c)=\sum_{i=0}^{N-1}[x(i)-\sum_{k=0}^{K}(i\Delta)^{k}c_{k}]^{2} \tag{1-7-2}$$

误差$E(c)$可以用微分方法使其达到最小，即将$E(c)$对c_j取偏导数，并令导数值为零而得到K+1个联立方程

$$\partial E/\partial c_{j}=\sum_{i=0}^{N-1}2[x(i)-\sum_{k=0}^{K}(i\Delta)^{k}c_{k}][-(i\Delta)^{j}]=0$$

$$j=0,\ 1,\ \cdots,\ K \tag{1-7-3}$$

整理之后得

$$\sum_{k=0}^{K}c_{k}\sum_{i=0}^{N-1}(i\Delta)^{k+j}=\sum_{i=0}^{N-1}(i\Delta)^{j}x(i) \tag{1-7-4}$$

$$j=0,\ 1,\ \cdots,\ K$$

当K较大时，求c很费事，但K大于3或4的情况很少，如果只要求提取低阶多项式，则可直接用系数算式来求。计算中通常取$\Delta=1$以提高计算精度。例如

K＝0时

$$c_0=\frac{1}{N}\sum_{i=0}^{N-1}x(i) \tag{1-7-5}$$

K＝1时

$$c_0=\frac{1}{N(N+1)}\left[2(2N-1)\sum_{i=0}^{N-1}x(i)-6\sum_{i=0}^{N-1}i\,x(i)\right] \tag{1-7-6}$$

$$c_1=\frac{12}{N^2-1}\left[\sum_{i=0}^{N-1}ix(i)-\frac{N-1}{2}\sum_{i=0}^{N-1}x(i)\right] \tag{1-7-7}$$

为简化计算可取数据个数N为奇数（若实际为偶数时可将最后一个数免去），并将多项式中i的变化范围改为对原点对称，即定义

$$\hat{x}(i)=\sum_{k=0}^{K}i^k d_k \qquad i=-[N/2],\ \cdots,\ [N/2]$$

下面将对i求和的符号简记为$\sum=\sum\limits_{i=-[N/2]}^{[N/2]}$，可求得

K＝1时

$$d_0=\Sigma x(i)/N,\qquad d_1=\Sigma ix(i)/\Sigma i^2$$

K＝2时

$$d_0=[\Sigma i^2\Sigma[i^2x(i)]-\Sigma x(i)\Sigma i^4]/[(\Sigma i^2)^2-N\Sigma i^4]$$

$$d_1=[\Sigma ix(i)]/\Sigma i^2$$

$$d_2=[\Sigma i^2\Sigma x(i)-N\Sigma i^2x(i)]/[(\Sigma i^2)^2-N\Sigma i^4]$$

K＝3时

$$d_0=[\Sigma x(i)\Sigma i^4-\Sigma i^2\Sigma i^2x(i)]/[N\Sigma i^4-(\Sigma i^2)^2]$$

$$d_1=[\Sigma i^4 \Sigma i^3 x(i)-\Sigma i^6 \Sigma i x(i)]/[(\Sigma i^4)^2-\Sigma i^2 \Sigma i^6]$$

$$d_2=[\Sigma i^2 \Sigma x(i)-N \Sigma i^2 x(i)]/[(\Sigma i^2)^2-N \Sigma i^4]$$

$$d_3=[\Sigma i^4 \Sigma i x(i)-\Sigma i^2 \Sigma i^3 x(i)]/[(\Sigma i^4)^2-\Sigma i^2 \Sigma i^6]$$

上式中

$\Sigma i^2=N(N^2-1)/12$，　$\Sigma i^4=N(N^2-1)(3N^2-7)/240$，

$\Sigma i^6=N(N^2-1)(3N^4-18N^2+31)/1344$

$\Sigma i^8=N(N^2-1)(5N^6-55N^4+239N^2-381)/11520$

$i=-\lfloor N/2 \rfloor, \cdots, \lfloor N/2 \rfloor$（$N$为奇数）

以上系数的计算可分两步进行：

（1）先计算矩量$\Sigma x(i)$，$\Sigma i^k x(i)$

（2）利用公式计算d_0，…，d_k

在计算中建议采用双精度。

尽管d_0，…d_k的计算形式比较方便，但d_0并不总是代表均值（除$K=0$和1外），为了让i从0开始，许多人宁愿用c_k而不用d_k，然而由d可以转换为c。为此先定义

$$M=-N\Delta/2$$

然后将i用$(M+i\Delta)/\Delta$代替（这时$i=0$，…，N），则

$$x(i)=\sum_{k=0}^{K}\left[\frac{M+i\Delta}{\Delta}\right]^k d_k \quad i=0, 1, \cdots, N$$

在$K=3$时

$$x(i)=d_0+\left(\frac{M+i\Delta}{\Delta}\right)d_1+\left(\frac{M+i\Delta}{\Delta}\right)^2 d_2+\left(\frac{M+i\Delta}{\Delta}\right)^3 d_3$$

$$=(d_0+(Md_1)/\Delta+(M^2 d_2)/\Delta^2+(M^3 d_3)/\Delta^3)$$
$$+i\Delta(d_1/\Delta+(2Md_2)/\Delta^2+(3M^2 d_3)/\Delta^3)$$
$$+(i\Delta)^2(d_2/\Delta^2+(3Md_3)/\Delta^3)+(i\Delta)^3 d_3/\Delta^3$$

使两边对应项相等可得

$$c_0=d_0+(Md_1)/\Delta+(M^2 d_2)/\Delta^2+(M^3 d_3)/\Delta^3$$

$$c_1=d_1/\Delta+(2Md_2)/\Delta^2+(3M^2 d_3)/\Delta^3$$

$c_2 = d_2/\Delta^2 + (3Md_3)/\Delta^3$

$c_3 = d_3/\Delta^3$

将适当的d_k项取为0，可以得到$K<3$的公式。附录二中的POLYTR程序就是根据上述公式编写的提取多项式趋势的程序。

1.8 随机数据的周期性检验

考察随机数据中是否有周期或近似周期分量，最直接的方法是分析产生数据的物理过程特性，由此甚至可以确定周期分量的频率。例如机械传动链的结构或对象其固有频率为已知的情况，或者一些具有明显季节性的现象等。但有些周期分量尚不明确或不明显的情况，它们和随机分量混杂在一起，需要加以检测或剔除。

从概率密度上讲，正弦信号的直方图和随机信号通常具有的钟形直方图迥然不同，它是盆形的（见图1-14）。但当正弦信号的方差比随机部分的方差小，或者包含具有一个以上的正弦信号时，就不容易从数据的直方图上直接判断周期分量的存在。

图1-14的图示为：$x(i)=\sin(0.15i)$，$i=0, 1, \cdots, 1000$，产生的数字正弦波，理论上应有

$$\mu=0, \quad \sigma^2=0.5$$

$$p(x)=\begin{cases} \dfrac{1}{\pi\sqrt{1-x^2}} & -1\leqslant x\leqslant 1 \\ 0 & |x|>1 \end{cases}$$

```
     SINE WAVE
NDATA =   1000
XBEG =   -1.000    XEND =     .920  DELTAX =     .080
 MEAN =     .000    STANDARD DEVIATION =     .707


 -1.040    .0      *
  -.960128.0   xxx*xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
  -.880 48.0   xxxx*xxxxxxx
  -.800 48.0   xxxxx*xxxxxx
  -.720 32.0   xxxxxx*x
  -.640 32.0   xxxxxx*x
  -.560 32.0   xxxxxxx*
  -.480 32.0   xxxxxxxx*
  -.400 32.0   xxxxxxxx *
  -.320 32.0   xxxxxxxx *
  -.240 16.0   xxxx       *
  -.160 32.0   xxxxxxxx   *
  -.080 16.0   xxxx       *
   .000 32.0   xxxxxxxx   *
   .080 32.0   xxxxxxxx   *
   .160 16.0   xxxx       *
   .240 32.0   xxxxxxxx   *
   .320 32.0   xxxxxxxx *
   .400 16.0   xxxx     *
   .480 32.0   xxxxxxxx*
   .560 32.0   xxxxxxx*
   .640 32.0   xxxxxx*x
   .720 32.0   xxxxxx*x
   .800 48.0   xxxxx*xxxxxx
   .880 48.0   xxxx*xxxxxxx
   .960136.0   xxx*xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
```

图1-14　正弦序列的直方图

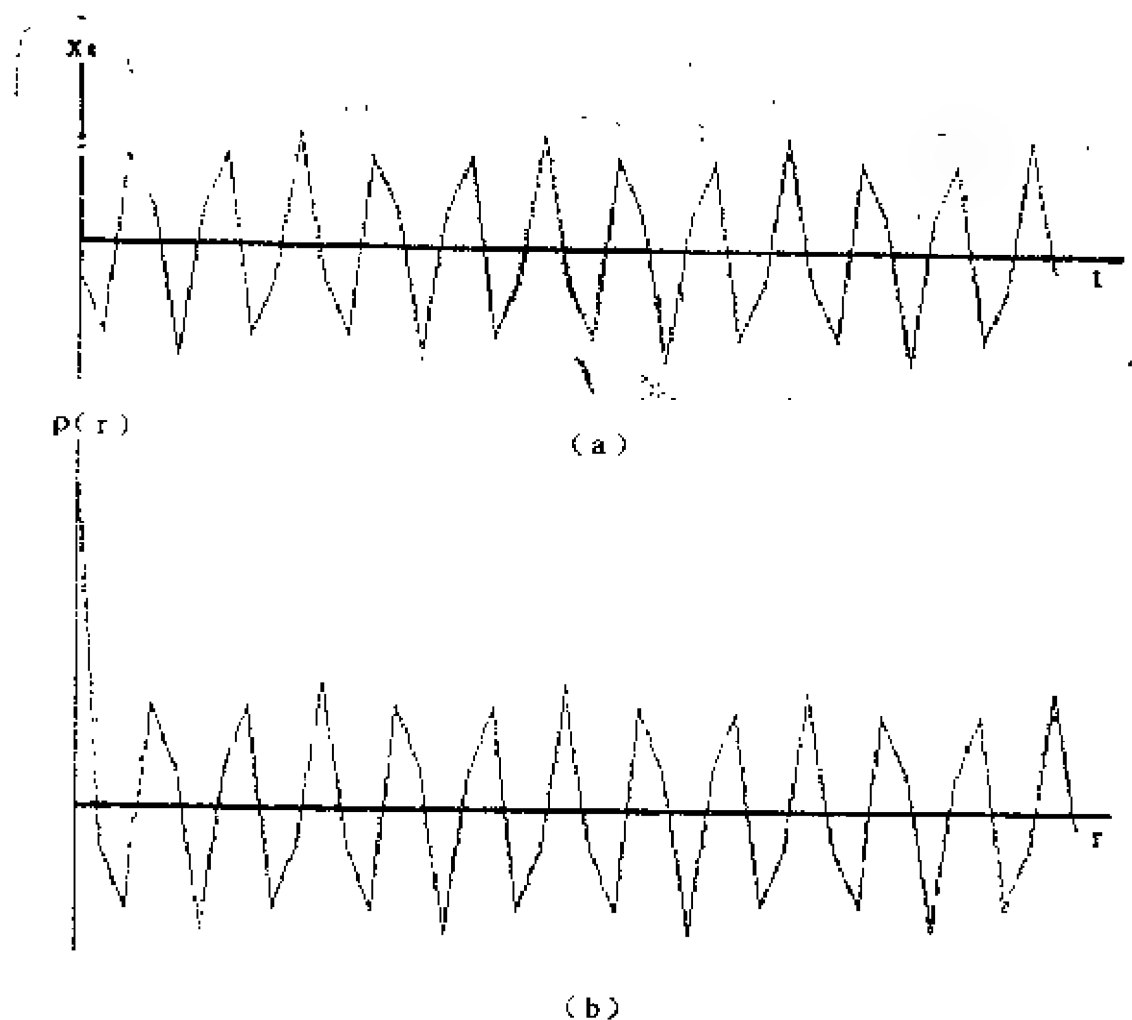
可以看出周期信号和高斯噪声不同，窄带高斯噪声不论带宽怎么窄，它们的概率密度总是高斯型的。这种噪声最可能的取值是在均值附近，而周期信号的概率密度都在均值处为最小。图1-15给出了正弦波叠加随机噪声的直方图。

```
0   SINE WAVE WITH NOISE
NDATA =   1000
 XBEG =  -1.000    XEND =      .920  DELTAX =      .080
 MEAN =    -.007   STANDARD DEVIATION =     .743

 -1.040  99.0   xxx*xxxxxxxxxxxxxxxxxx
  -.960  45.0   xxxx*xxxxxx
  -.880  37.0   xxxx*xxxx
  -.800  45.0   xxxxx*xxxxx
  -.720  37.0   xxxxxx*xx
  -.640  37.0   xxxxxx*xx
  -.560  33.0   xxxxxxx*
  -.480  25.0   xxxxxx   *
  -.400  28.0   xxxxxxx *
  -.320  35.0   xxxxxxxxx*
  -.240  21.0   xxxxx     *
  -.160  27.0   xxxxxxx     *
  -.080  31.0   xxxxxxxxx  *
   .000  29.0   xxxxxxx     *
   .080  32.0   xxxxxxxx    *
   .160  23.0   xxxxxx     *
   .240  29.0   xxxxx      *
   .320  17.0   xxxx       *
   .400  33.0   xxxxxxxx*
   .480  25.0   xxxxxx   *
   .560  35.0   xxxxxxx*x
   .640  44.0   xxxxxx*xxxx
   .720  39.0   xxxxxx*xxx
   .800  41.0   xxxxx*xxxx
   .880  43.0   xxxx*xxxxxx
   .960 124.0   xxxx*xxxxxxxxxxxxxxxxxxxxxxxxx
```

图1-15　正弦波叠加噪声的直方图

从自相关函数看，当迟后较大时，会有一个或几个周期分量的相关函数曲线是不衰减的振荡，而且振荡的频率也就是周期分量的频率。图1-16 a，b，分别为一周期序列（带有2%振幅的噪声）及其自相关函数图形。



1-16　周期序列(a)及其自相关(b)

在频率域内分析周期信号是应用极广的，第四章中的周期图方法将给出有关这方面的较为详尽的论述，因此本节不再作进一步的讨论。

第二章　平稳随机过程及其模型

2.1　平稳随机过程

随机过程的类型很多，至今还没有适合各种类型随机过程的分析方法，这里只讨论称为“平稳的”和可以平稳化的特殊的非平稳随机过程。

平稳过程的特点是它的统计特性不随时间而变化，这通常是当一个稳定的物理系统达到稳态时所产生的过程。这种状态有时也称为“统计平衡”状态。如果 $\{x(t)\}$ 是这种过程，因其统计特性不随时间改变，则$x(1)$，$x(2)\cdots$，$x(t)$，$\cdots$都应有同样的一维概率密度(PDF)，但统计特性不只是这些，还应包括 $\{x(1), x(4)\}$，$\{x(2), x(5)\}$，$\{x(3), x(6)\}$，$\cdots$具有相同的二维PDF，进一步还有 $\{x(1), x(3), x(6)\}$，$\{x(2), x(4), x(7)\}$，$\cdots$等具有相同的三维PDF，如此等等．总之，对时间点t_1，t_2，$\cdots$，t_n的任何集合，$\{x(t_1), x(t_2), \cdots, x(t_n)\}$ 的联合概率分布在每个时间点移动同一数值时均保持不变。如果 $\{x(t)\}$ 具备这种性质，则称其为“完全平稳”的、换言之，完全平稳随机过程的概率结构对时间原点的移动保持不变。

完全平稳的要求是苛刻的，因此引入一种较弱的条件——“m阶平稳”以图概略地描述同样的物理过程。这种条件不要求例如$x(t_1)$的概率分布必须和$x(t_1+k)$的概率分布相同，而仅要求这两个分布的主要特征相同，譬如只要求相等到某阶矩；同样，不坚持 $\{x(t_1),x(t_2)\}$ 的联合分布和 $\{x(t_1+k), x(t_2+k)\}$ 的联合分布等同，而只要求到某阶联合矩相等。依此类推，由这个基本思想引出如下定义：

如果对任何$t_1,t_2,\cdots,t_n$和k，$\{x(t_1),x(t_2)\cdots,x(t_n)\}$的直到$m$

阶的联合矩存在，并且和$\{x(t_1+k),x(t_2+k),\cdots,x(t_n+k)\}$的相应的各阶联合矩相等，则称$\{x(t)\}$为$m$阶平稳过程。用等式表达如下

$$E[(x(t_1))^{m_1}(x(t_2))^{m_2}\cdots(x(t_n))^{m_n}]$$
$$=E[(x(t_1+k))^{m_1}(x(t_2+k))^{m_2}\cdots(x(t_n+k))^{m_n}] \quad (2-1-1)$$

其中k任意，m_1，…，m_n为任意正整数，且$m_1+m_2+\cdots+m_n\leqslant m$。

式(2-1-1)是m阶平稳的一般要求，作为其中的一个要求，例如除m_1外其他$m_2=m_3=\cdots=m_n=0$，取$k=-t$，则对任何t和所有$m_1\leqslant m$有

$$E[(x(t))^{m_1}]=E[(x(0))^{m_1}]=\text{常数(与}t\text{无关)}$$

若m_1和m_2非零，而$m_3=m_4\cdots=m_n=0$，则对任何t，s和所有满足$m_1+m_2\leqslant m$的m_1和m_2有

$$E[(x(t))^{m_1}(x(s))^{m_2}]=E[(x(0))^{m_1}(x(s-t))^{m_2}]$$
$$=\text{只和}(s-t)\text{有关的函数}$$

对于平稳阶数最常用的是

(1)一阶平稳($m=1$)——如果$\{x(t)\}$是一阶平稳的，则意味着$E[x(t)]=\mu$是和t无关的常数。

(2)二阶平稳($m=2$)——如果$\{x(t)\}$是二阶平稳的，则以下三种可能的情况均需满足，即

Ⅰ. $E[x(t)]=\mu$是和t无关的常数。

Ⅱ. $E[(x(t))^2]=\mu_2$是和t无关的常数。

因而有$\mathrm{Var}[x(t)]=\mu_2-\mu^2=\sigma^2$也是和$t$无关。

Ⅲ. 对任何t，s，$E[x(t)x(s)]$是且只是$(t-s)$的函数。因而有

$$\mathrm{Cov}[x(t),\ x(s)]=E[x(t)x(s)]-\mu^2=\text{只是}(t-s)\text{的函数}$$

因此二阶平稳过程的性质归结为：对所有的时间点，过程具有同样的均值(μ)、方差(σ^2)，而且任何二时间点s，t之间过程的协方差只取决于时间间隔$(t-s)$，而和这些点在时间轴上的位

置无关。

我们知道，在一般情况下，矩的无限序列唯一地确定了概率分布，给定的矩愈多，就能愈准确地给定分布。自然地希望有一个“转折点”m_0，当过程为m_0阶或更高阶平稳时，从实用上看该过程就近乎完全平稳了。通常作为表征过程功率的函数是过程的二次函数，即过程的功率特性只取决于二阶特性，如果过程是高斯的，则二阶平稳必然意味着完全平稳。以后我们把至少为二阶平稳的过程称为平稳过程，它更准确的名称是弱平稳或广义平稳。

2.2 自协方差和自相关函数

由平稳随机过程的性质引出的两个重要函数是自协方差函数和自相关函数。

如果$x(t)$定义于t的离散值$t=0$，± 1，$\pm 2\cdots$并把这种离散过程记作$\{x_t\}$，则

$$\begin{aligned}\operatorname{Cov}[x_t,\ x_{t+r}]&=E[(x_t-\mu)(x_{t+r}-\mu)]\\&\equiv R(r)\qquad r=0,\ \pm 1,\ \pm 2,\ \cdots\end{aligned}\tag{2-2-1}$$

$R(r)$只是r的函数，称为$\{x_t\}$的自协方差函数。对每个r，$R(r)$是过程在相隔时间为r的一对值之间的协方差，称r为“迟后”。

自相关函数定义为

$$\rho(r)\equiv R(r)/R(0)\qquad r=0,\ \pm 1,\ \pm 2,\ \cdots\tag{2-2-2}$$

由于

$$R(0)=E[(x_t-\mu)^2]=\operatorname{Var}[x_t]=\sigma^2\tag{2-2-3}$$

故

$$\rho(r)=\frac{\operatorname{Cov}[x_t,\ x_{t+r}]}{[\operatorname{Var}(x_t)\cdot\operatorname{Var}(x_{t+r})]^{1/2}}\tag{2-2-4}$$

对每个r，$\rho(r)$表示在相隔时间为r的一对值之间的相关系数，它可以作为x_t的某一实现和时移r后的同一实现之间"相似"的量度。

自协方差函数具有如下性质：

(1)$R(0)=\sigma^2$

(2)$|R(r)|\leqslant R(0)$　　对所有r

(3)若$\{x_t\}$为实数序列，则$R(-r)=R(r)$，即$R(r)$为偶函数，因为

$$R(-r)=E[(x_t-\mu)(x_{t-r}-\mu)]=E[(x_{t+r}-\mu)(x_t-\mu)]$$
$$=R(r)$$

类似地，自相关函数有如下性质

(1)$\rho(0)=1$

(2)$|\rho(r)|\leqslant 1$　　对所有r

(3)$\rho(-r)=\rho(r)$　　对所有r(x_t为实序列)

可见，$R(r)$和$\rho(r)$均对称于原点($r=0$)，且在原点处具有最大值。

前面假定对每一时间点，过程的值均取为实数，实际中有时用复数表示过程更方便（例如在电路中常用复数来表示电压），设$\{x_t\}$为复数序列，即

$$x_t=u_t+jv_t$$

其中$j=\sqrt{-1}$，而u_t和v_t均为实数序列，则$\{x_t\}$的均值定义为

$$E[x_t]=E[u_t]+jE[v_t]=\mu(\text{与}t\text{无关})$$

$\{x_t\}$的自协方差定义为

$$R(r)=E[(x_t^*-\mu^*)(x_{t+r}-\mu)] \tag{2-2-5}$$

其中*表示复共轭。复序列的$R(r)$一般也是复数，但

$$\mathrm{Var}[x_t]=R(0)=E[|x_t-\mu|^2]$$

总是实数，类似地，自相关函数定义为

$$\rho(r)=R(r)/R(0) \tag{2-2-6}$$

它一般也是复数。对于复序列，性质(3)应改为

$$R(-r)=R^*(r) \tag{2-2-7}$$

$$\rho(-r)=\rho^*(r) \tag{2-2-8}$$

自相关函数另一个有用的性质是：若定义一个时间序列(x_0, x_1, x_2, …, x_N)的逆序列为(x_N^*, x_{N-1}^*, …, x_1^*, x_0^*)，那么可以证明这个逆序列的自相关函数和原序列的自相关函数完全相同。对于实数序列，其逆序列就是(x_N, x_{N-1}, …, x_1, x_0)，图2-1是逆序列和原序列的对照。

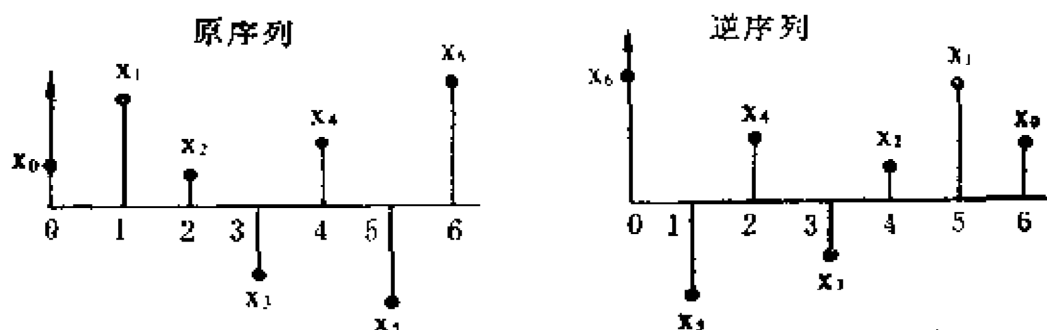


图2-1 原序列和它的逆序列

自协方差函数或自相关函数还具有一个重要性质，这就是它所构成的协方差阵R是非负定的，或者说行列式$\det R\geqslant 0$，这里

$$R=\begin{pmatrix} R(0) & R(1)\cdots & R(N) \\ R(-1) & R(0)\cdots & R(N-1) \\ \cdots\cdots & & \\ R(-N) & R(-N+1)\cdots & R(0) \end{pmatrix} \tag{2-2-9}$$

为证明这一性质，先写出矩阵特征方程

$$R\mathbf{e}=\lambda\mathbf{e} \tag{2-2-10}$$

其中λ是R的特征值，$\mathbf{e}$是相应的特征向量。在式(2-2-10)两边前乘$\mathbf{e}$的共轭转置$\mathbf{e}^H$，

$$\mathbf{e}^H R\mathbf{e}=\lambda\mathbf{e}^H\mathbf{e}$$

由于

$$\mathbf{e}^H R\mathbf{e}=\sum_{i=0}^{N}\sum_{j=0}^{N}e_i^* R(i-j)e_j=\sum_{i=0}^{N}\sum_{j=0}^{N}e_i^* E\ \{x_i x_j^*\}\ e_j$$

$$=E\left\{\left(\sum_{i=0}^{N}e_i^* x_i\right)\left(\sum_{j=0}^{N}e_j x_j^*\right)\right\}=E\left\{\left|\sum_{i=0}^{N}e_i^* x_i\right|^2\right\}\geqslant 0$$

且

$$\mathbf{e}^H\mathbf{e}=\sum_{i=0}^{N}e_i^* e_i=\sum_{i=0}^{N}|e_i|^2>0$$

故

$$\lambda=\mathbf{e}^H R\mathbf{e}/(\mathbf{e}^H\mathbf{e})\geqslant 0 \qquad (2-2-11)$$

即R的特征值为非负的实数。

以R的特征向量$\mathbf{e}_0$，$\mathbf{e}_1$，…，$\mathbf{e}_N$为列构成$(N+1)\times(N+1)$矩阵$E\equiv(\mathbf{e}_0,\ \mathbf{e}_1,\ \cdots,\ \mathbf{e}_N)$；以$R$的特征值为元素构成对角阵

$$\mathbf{\Lambda}=\begin{pmatrix}\lambda_0 & & O\\ & \lambda_1 & \\ & & \ddots & \\ O & & & \lambda_N\end{pmatrix}$$

于是有

$$RE=E\mathbf{\Lambda}$$

这实际是集中了N+1个特征向量的特征方程，对它两边求行列式的值，则有

$$\det R\cdot\det E=\det E\cdot\det\mathbf{\Lambda}$$

或 $$\det R=\det\mathbf{\Lambda}=\prod_{i=0}^{N}\lambda_i\geqslant 0$$

最后一个不等式是根据式(2-2-11)得出的，这就证明了协方差阵R是非负定的性质。

对于由自相关函数$\rho(r)$构成的相关阵也可以得出同样的结

论。这个性质表明，不能任意选择函数作为$R(r)$或$\rho(r)$，即便它满足性质(1)—(3)。能够代表某种过程自协方差或自相关的函数必须是非负定的。

2.3 典型离散参数模型

在讨论实际数据的模型拟合之前，先看一些典型离散过程所对应的模型。

1.纯随机过程——白噪声 随机过程$x_t(t=0, \pm1, \pm2, \cdots)$如果是由一个无关的随机变量的序列构成，即对于所有$s\neq t$，$\mathrm{Cov}(x_s, x_t)=0$，则称其为纯粹随机过程。这是所有离散参数模型中最简单的模型，它相当于没有“记忆”的过程，即t时刻过程的值和所有过去的直到$(t-1)$时刻的值（实际也包括过程的未来值）都不相关。图2-2表示由零均值单位方差的正态纯随机过程得到的一个实现（200个观察值）。

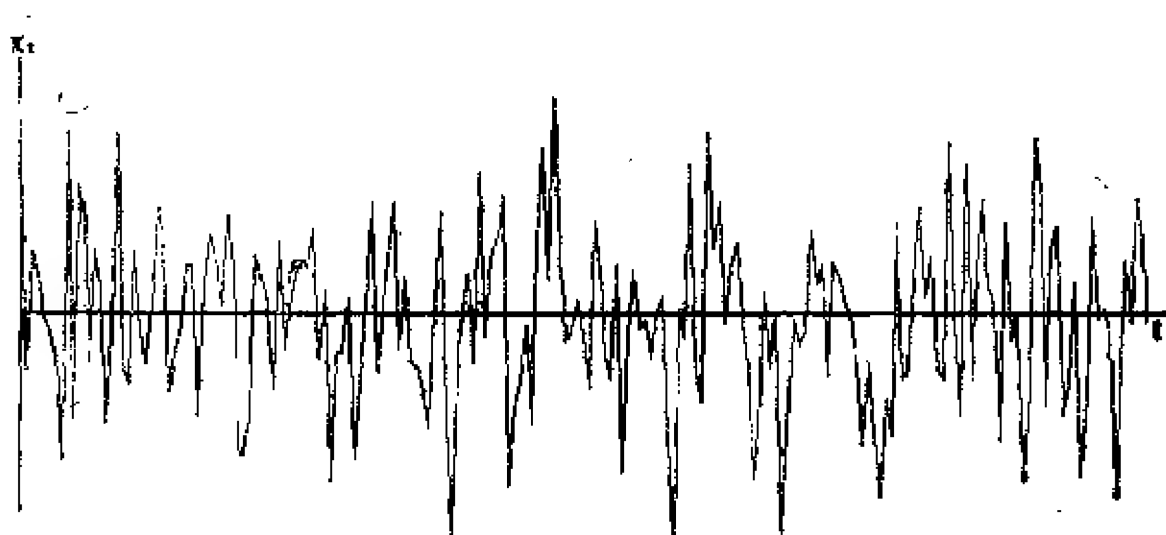


图2-2 一个正态纯随机过程的实现

对于这类过程的平稳性要求是

$$E(x_t)=\mu_x=\text{const} \quad 对所有t$$

和

$$E((x_t-\mu)^2)=\sigma_x^2=\text{const} \quad 对所有t$$

于是自协方差函数为

$$R(r)=\text{Cov}(x_t, x_{t+r})=\begin{cases}\sigma_x^2 & r=0\\ 0 & r\neq 0\end{cases} \tag{2-3-1}$$

自相关函数为(见图2-3)

$$\rho(r)=\begin{cases}1 & r=0\\ 0 & r\neq 0\end{cases} \tag{2-3-2}$$

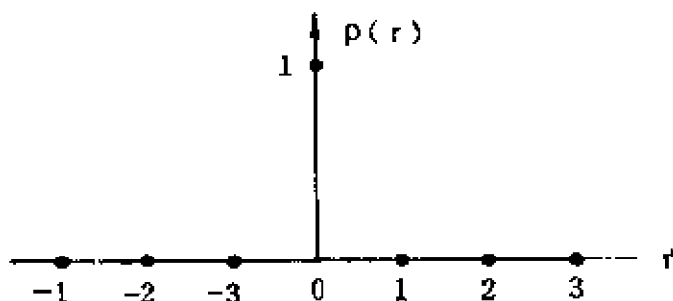


图2-3 纯随机过程（白噪声）的自相关

以后提到纯随机过程时都是指这种平稳的纯随机过程，它也称为"白噪声"。

应当指出，纯随机过程模型是人为的，在实际中过程前后往往都存在"记忆"，然而这种模型为构造更为复杂的模型提供了基本"元素"，因此它在平稳过程理论的发展中起着重要的作用。

2.一阶自回归过程 比纯随机过程复杂一点的是"一步相关"模型，亦即过程在t时刻的值至少部分地和$(t-1)$时刻的值相关，这种模型称为一阶自回归模型，定义如下：

若 $\{x_t\}$ 满足差分方程

$$x_t + a x_{t-1} = \varepsilon_t \tag{2-3-3}$$

其中a为常数，$\{\varepsilon_t\}$ 为（平稳）纯随机过程，则称 $\{x_t\}$ 为一阶自回归过程，并记为$AR(1)$。式(2-3-3)也可改写成

$$x_t = -a x_{t-1} + \varepsilon_t \tag{2-3-4}$$

可见x_t部分地依赖于x_{t-1}，部分地依赖于“随机扰动”ε_t，或者说x_t“线性回归”到x_{t-1}，而ε_t为“误差”项。由于x_t具有对自身过去的回归，因而称之为“自回归过程”。图2-4为$AR(1)$模型$x_t - 0.65\, x_{t-1} = \varepsilon_t$（$\varepsilon_t$为零均值单位方差的正态随机过程）的某一实现的200个观察值。

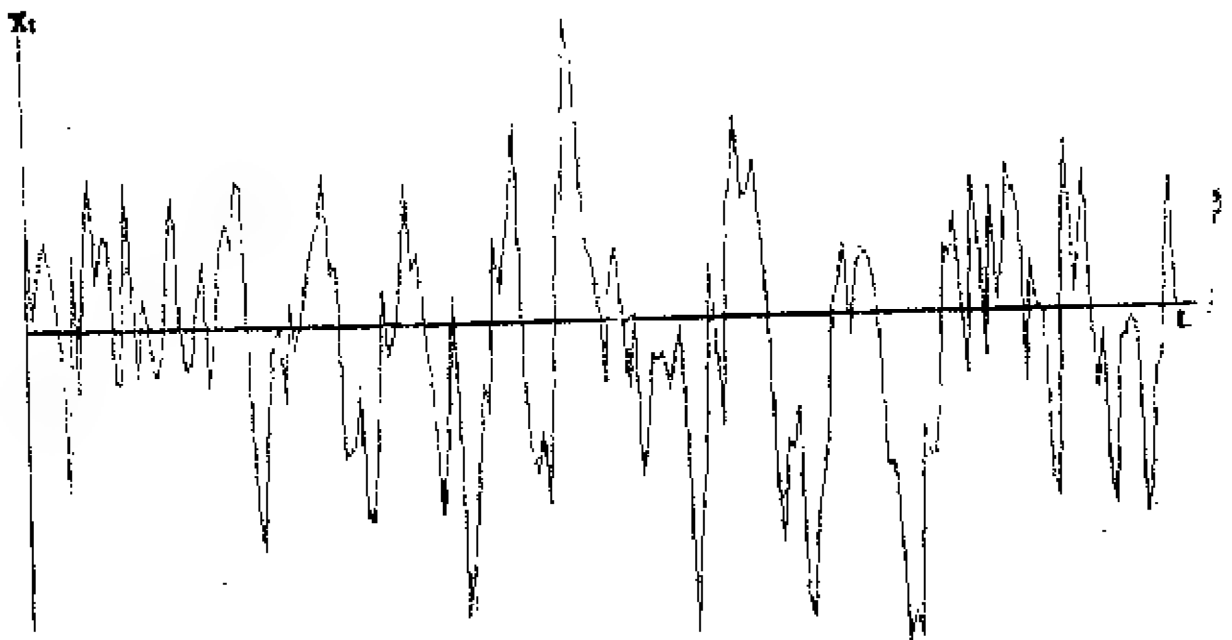


图2-4 一个AR (1) 过程的实现

在分析该模型的二阶统计特性之前，先求解上述差分方程，即用ε_t，ε_{t-1}，…来表达x_t。为此设初始条件$x_0 = 0$(即$x_1 = \varepsilon_1$)。用迭代法可以解得

$$x_t = \varepsilon_t - a x_{t-1} = \varepsilon_t - a(\varepsilon_{t-1} - a x_{t-2}) = \cdots$$

$$= \varepsilon_t - a\varepsilon_{t-1} + a^2\varepsilon_{t-2} - \cdots + (-a)^{t-1}\varepsilon_1 \tag{2-3-5}$$

若$E(\varepsilon_t) = \mu_\varepsilon$(所有$t$)，则

$$E(x_t)=\mu_\varepsilon(1-a+a^2-\cdots+(-a)^{t-1})$$

$$=\begin{cases}\mu_\varepsilon t & a=-1\\ \mu_\varepsilon\left(\dfrac{1-(-a)^t}{1+a}\right. & a\neq-1\end{cases} \tag{2-3-6}$$

如果$\mu_\varepsilon=0$，则$E(x_t)=0$(所有t)，$\{x_t\}$是一阶平稳的。如果$\mu_\varepsilon\neq0$，则$E(x_t)$是t的函数，因此$\{x_t\}$连一阶平稳也达不到，然而如果$|a|<1$，则当t很大时

$$E(x_t)\simeq\mu_\varepsilon/(1+a) \tag{2-3-7}$$

称$\{x_t\}$是一阶渐近平稳的

后面假定$\mu_\varepsilon=0$，于是

$$\mathrm{Var}(\varepsilon_t)=E(\varepsilon_t^2)=\sigma_\varepsilon^2 \tag{2-3-8}$$

$$\mathrm{Cov}(\varepsilon_t,\ ,\ \varepsilon_s)=E(\varepsilon_t\varepsilon_s)=0\quad s\neq t \tag{2-3-9}$$

现在看$\{x_t\}$的方差和协方差

$$\sigma_x^2=\mathrm{Var}(x_t)=E[(\varepsilon_t-a\varepsilon_{t-1}+\cdots+(-a)^{t-1}\varepsilon_t)^2]$$

$$=\sigma_\varepsilon^2(1+a^2+a^4+\cdots a^{2t-2})$$

$$=\begin{cases}\sigma_\varepsilon^2(1-a^{2t})/(1-a^2) & |a|\neq1\\ \sigma_\varepsilon^2 t & |a|=1\end{cases} \tag{2-3-10}$$

$$\mathrm{Cov}(x_t,\ x_{t+\tau})=E(x_t x_{t+\tau})$$

$$=E[(\varepsilon_{t+\tau}-a\varepsilon_{t+\tau-1}+\cdots+(-a)^{t+\tau-1}\varepsilon_1)$$

$$\cdot(\varepsilon_t-a\varepsilon_{t-1}+\cdots+(-a)^{t-1}\varepsilon_1)]$$

$$=\sigma_\varepsilon^2[(-a)^\tau+(-a)^{\tau+2}+\cdots+(-a)^{\tau+2(t-1)}]$$

$$=\begin{cases}\sigma_\varepsilon^2(-a)^\tau(1-a^{2t})/(1-a^2) & |a|\neq1\\ \sigma_\varepsilon^2 t & a=-1\\ \sigma_\varepsilon^2(-1)^\tau t & a=1\end{cases} \tag{2-3-11}$$

可见$\mathrm{Var}(x_t)$和$\mathrm{Cov}(x_t,\ x_{t+\tau})$都是$t$的函数，因此$\{x_t\}$不是二阶平稳的，但若$|a|<1$，则当$t$足够大时

$$\sigma_x^2 \simeq \sigma_\varepsilon^2/(1-a^2) \tag{2-3-12}$$

$$\mathrm{Cov}[x_t,\ x_{t+r}] \simeq \sigma_\varepsilon^2(-a)^r/(1-a^2) \tag{2-3-13}$$

故 $\{x_t\}$ 是二阶渐近平稳的。注意这种二阶渐近平稳的条件是 $|a|<1$。

乍看起来，似乎任何过程都会是渐近平稳的，因为任何情况 $\lim\limits_{t\to\infty}\mathrm{Var}[x_t]$ 都和 t 无关，而 $\lim\limits_{t\to\infty}\mathrm{Cov}[x_t, x_{t+r}]$ 只是 r 的函数，但问题的关键是这些极限是否存在，即当 $t\to\infty$ 时 $\mathrm{Var}(x_t)$ 和 $\mathrm{Cov}[x_t, x_{t+r}]$ 可能并不收敛到有限的值。例如 $a=-1$ 时的 $AR(1)$ 模型，由式(2-3-10)可见其 $\mathrm{Var}[x_t]$ 随 t 线性增大，当 $t\to\infty$ 时 $\mathrm{Var}[x_t]\to\infty$，这时对应的式(2-3-5)为

$$x_t=\sum_{s=1}^{t}\varepsilon_s \tag{2-3-14}$$

它称为“随机游动”模型

现在还是回到满足平稳条件 $|a|<1$ 的情况，这时的自协方差函数为

$$R(r)=\sigma_\varepsilon^2(-a)^{|r|}/(1-a^2) \qquad r=0,\pm1,\pm2,\cdots \tag{2-3-15}$$

而自相关函数为

$$\rho(r)=R(r)/R(0)=(-a)^{|r|} \quad r=0,\ \pm1,\ \pm2\cdots \tag{2-3-16}$$

图2-5画出 $a>0$ 和 $a<0$ 时 $\rho(r)$ 的图形，它们的衰减均受指数控制。

在前面讨论中，由于给定初始条件($x_0=0$)，x_t 不是平稳的，但当 $t\to\infty$ 时可以成为平稳，即“渐近平稳”。事实上，由于我们规定在 $t=0$ 时过程处于特定状态（即 x_0 为零，过程的所有实现都要通过原点），因此对时间原点移动的不变性这一特点不存在了，只有当过程“忘掉”初始值之后才进入平稳状态。这种情况类似于线性电路 $t=0$ 时开关接通后输出端的过渡过程。

如果用常系数线性差分方程的一般方法解式(2-3-3)，则解由

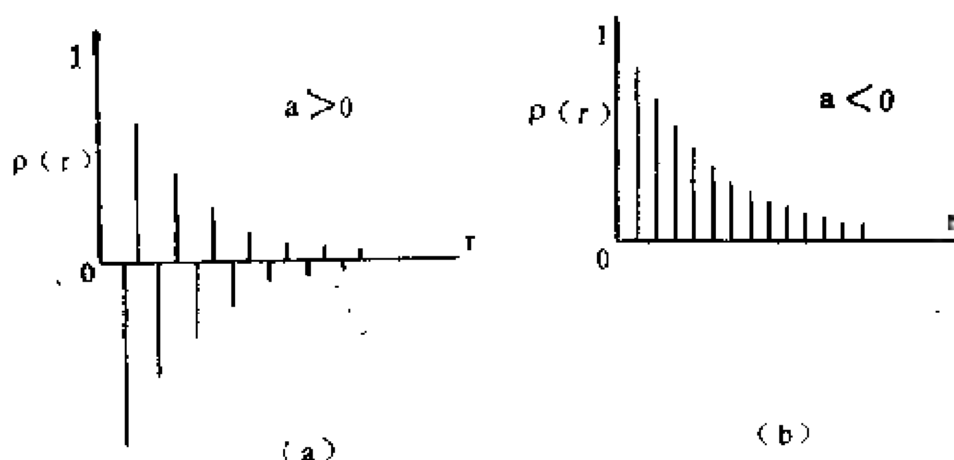


图2-5 AR (1) 过程的自相关

两部分组成——余函数和特解。在$AR(1)$的情况下，余函数是齐次方程$x_t+ax_{t-1}=0$的解$A(-a)^t$（A为任意常数），特解用"后移算子"B最容易求，B定义为后移一步的运算即

$$Bx_t=x_{t-1} \tag{2-3-17}$$

利用它将(2-3-3)改写成

$$(1+aB)x_t=\varepsilon_t \tag{2-3-18}$$

可得特解为

$$x_t=(1+aB)^{-1}\varepsilon_t=\left[\sum_{s=0}^{\infty}(-a)^sB^s\right]\varepsilon_t$$

$$=\sum_{s=0}^{\infty}(-a)^s\varepsilon_{t-s} \tag{2-3-19}$$

于是式(2-3-3)的通解为

$$x_t=A(-a)^t+\sum_{s=0}^{\infty}(-a)^s\varepsilon_{t-s} \tag{2-3-20}$$

A由初始条件决定，如果$x_0=0$，则$A=-\sum_{s=0}^{\infty}(-a)^s\varepsilon_{-s}$，而式

(2-3-20)可化作式(2-3-5)。特解项是无穷多个均值为零、方差有限(σ_ε^2)的不相关随机变量之和。当$|a|<1$的情况下

$$\mathrm{Var}\left[\sum_{s=0}^{\infty}(-a)^s\varepsilon_{t-s}\right]=\sigma_\varepsilon^2\sum_{s=0}^{\infty}a^{2s}=\sigma_\varepsilon^2/(1-a^2)<\infty$$

显然，在$|a|<1$时，由于余函数在$t\to\infty$时衰减为零，因此式(2-3-19)完全确定了"稳态"时$\{x_t\}$的行为，所以这个特解也称为式(2-3-3)的"平稳解"。

如果事先假定$\{x_t\}$是平稳的，或者说$\{x_t\}$已进入平稳，则可以用另一种方法推导其自协方差函数，这就是在式(2-3-4)两边乘以$x_{t-r}(r\geqslant1)$，然后取数学期望，即

$$E[x_tx_{t-r}]=-aE[x_{t-1}x_{t-r}]+E[\varepsilon_tx_{t-r}] \tag{2-3-21}$$

由于x_{t-r}是ε_{t-r}，ε_{t-r-1}，…的线性函数(见式(2-3-5))，当$r\geqslant1$时它和ε_t是无关的，因此$E[\varepsilon_tx_{t-r}]=0$，于是式(2-3-21)成为

$$R(r)=-aR(r-1)\qquad r\geqslant1 \tag{2-3-22}$$

同样有

$$\rho(r)=-a\rho(r-1)=a^2\rho(r-2)\cdots=(-a)^r\rho(0)$$
$$=(-a)^r \tag{2-3-23}$$

考虑到$\rho(r)$是r的偶函数，故

$$\rho(r)=(-a)^{|r|}\qquad r=0,\ \pm1,\ \pm2,\ \cdots \tag{2-3-24}$$

这和式(2-3-16)所得相同。同样地，在式(2-3-4)两边乘以x_t，可得

$$\sigma_x^2=E[x_t^2]=-aE[x_{t-1}x_t]+E[\varepsilon_tx_t]$$

因$x_t=\varepsilon_t+(\varepsilon_{t-1},\varepsilon_{t-2},\cdots$的线性函数)，故$E[\varepsilon_tx_t]=E[\varepsilon_t^2]=\sigma_\varepsilon^2$。利用式(2-3-22)当$r=1$时$R(1)=-a\sigma_x^2=E[x_{t-1}x_t]$，可以得到和式(2-3-12)同样的结果。

3.二阶自回归过程 从$AR(1)$进一步扩展自然会引出"两步

相关"的模型，即x_t既与x_{t-1}又与x_{t-2}线性相关。定义如下：若过程$\{x_t\}$满足差分方程

$$x_t+a_1x_{t-1}+a_2x_{t-2}=\varepsilon_t \tag{2-3-25}$$

其中a_1，a_2为常数$\{\varepsilon_t\}$为纯随机过程，$E(\varepsilon_t)=0, E(\varepsilon_t^2)=\sigma_\varepsilon^2$，则称$\{x_t\}$为二阶自回归过程并记作$AR(2)$。图2-6为模型$x_t-0.45x_{t-1}+0.75x_{t-2}=\varepsilon_t$($\varepsilon_t$为零均值单位方差的正态随机过程)所产生的一个实现(200个观察值)的记录。

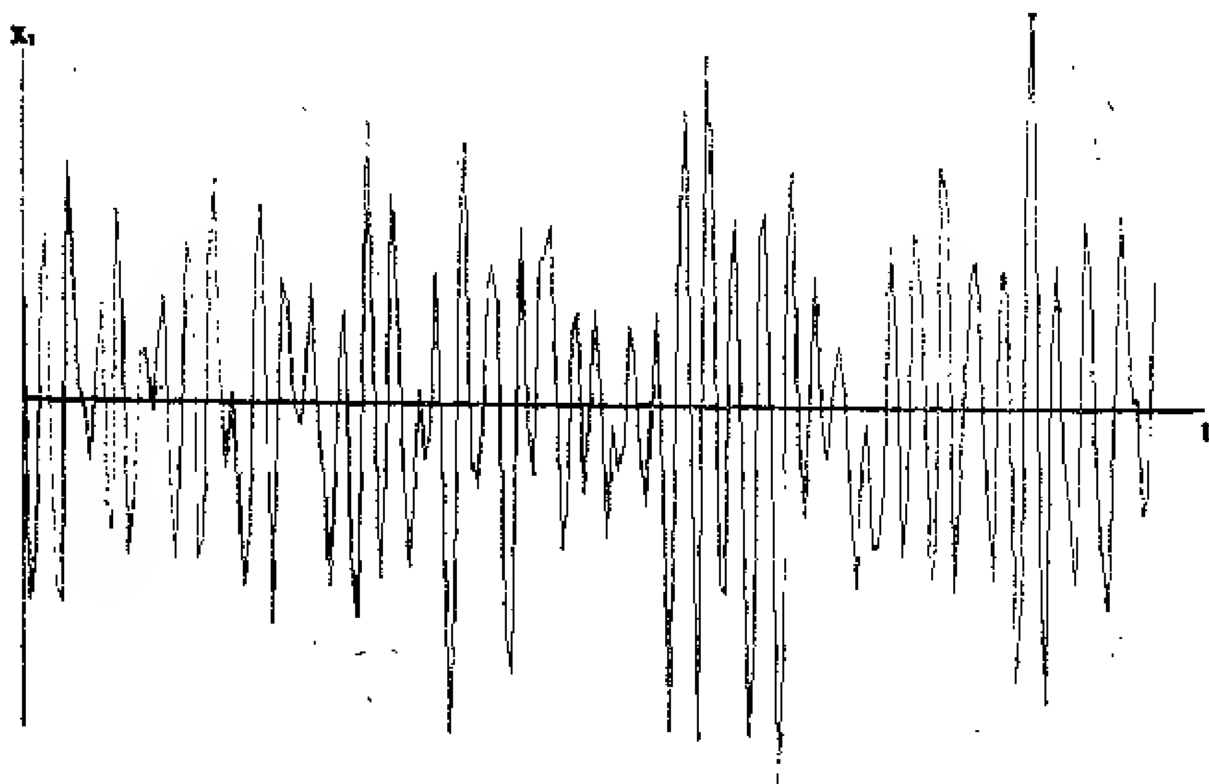


图2-6　一个AR（2）过程的实现

下面通过差分方程式(2-3-25)的通解来分析渐近平稳条件，利用后移算子将上式改写为

$$(1+a_1B+a_2B^2)x_t=\varepsilon_t \tag{2-3-26}$$

或

$$(1-\lambda_1B)(1-\lambda_2B)x_t=\varepsilon_t \tag{2-3-27}$$

其中λ_1，λ_2（假设不等）为$f(z)=z^2+a_1z+a_2$的根，即

$$\lambda_{1,2}=(-a_1\pm\sqrt{a_1^2-4a_2})/2 \tag{2-3-28}$$

式(2-3-26)的特解为

$$x_t=\frac{1}{(1-\lambda_1 B)(1-\lambda_2 B)}\varepsilon_t=\frac{1}{\lambda_1-\lambda_2}\left[\frac{\lambda_1}{1-\lambda_1 B}-\frac{\lambda_2}{1-\lambda_2 B}\right]\varepsilon_t$$

$$=\frac{1}{\lambda_1-\lambda_2}\left[\sum_{s=0}^{\infty}(\lambda_1^{s+1}-\lambda_2^{s+1})B^s\right]\varepsilon_t$$

$$=\sum_{s=0}^{\infty}\left(\frac{\lambda_1^{s+1}-\lambda_2^{s+1}}{\lambda_1-\lambda_2}\right)\varepsilon_{t-s} \tag{2-3-29}$$

余函数是齐次方程$x_t+a_1x_{t-1}+a_2x_{t-2}=0$的解：

$$x_t=A_1\lambda_1^t+A_2\lambda_2^t \tag{2-3-30}$$

A_1，A_2为任意常数，故通解为

$$x_t=A_1\lambda_1^t+A_2\lambda_2^t+\sum_{s=0}^{\infty}\left(\frac{\lambda_1^{s+1}-\lambda_2^{s+1}}{\lambda_1-\lambda_2}\right)\varepsilon_{t-s} \tag{2-3-31}$$

和$AR(1)$的讨论类似，渐近平稳的条件为

$$|\lambda_1|<1 \quad 和 \quad |\lambda_2|<1 \tag{2-3-32}$$

条件满足时，特解部分具有有限的方差，从而在均方意义下收敛，即式(2-3-29)代表平稳解。不难证明，上述条件也保证当$t\to\infty$时，$E[x_t^2]$和$E[x_tx_{t+\tau}]$都具有有限的极限。

通常对于$AR(2)$模型的平稳条件是直接用系数a_1，a_2的允许域来表达。它的确定，自然是从式(2-3-32)出发，同时还利用了关系式

$$\lambda_1+\lambda_2=-a_1 \quad 和 \quad \lambda_1\lambda_2=a_2$$

显然，$|a_2|<1$是系数应满足的条件之一。

在$\lambda_1<1$不等式两边乘以$(1-\lambda_2)$(它应为正)可得$\lambda_1+\lambda_2-\lambda_1\lambda_2<1$，即$a_1+a_2>-1$。在$\lambda_1>-1$不等式两边乘以$(1+\lambda_2)$(它亦应为正)可得$\lambda_1+\lambda_2+\lambda_1\lambda_2>-1$，即$a_1-a_2<1$。

上述条件综合起来，就是要求a_1，a_2的取值落在图2-7中由下式确定的三角形区域内

$$a_2<1,\ a_1+a_2>-1,\ a_1-a_2<1 \tag{2-3-33}$$

这个条件也可以用其它方法（如后面提到的朱利准则）加以证明。

在三角形系数域内还可以分成三种情况：

(1)临界阻尼——系数关系满足$a_2=a_1^2/4$(图2-7中抛物线AOB)，此时$\lambda_1=\lambda_2$，即模型有相等的实根。

(2)欠阻尼——$a_2>a_1^2/4$，$\lambda_1=\lambda_2^*$即二共轭复根。

(3)过阻尼——$a_2<a_1^2/4$，λ_1和λ_2为不等的实根。

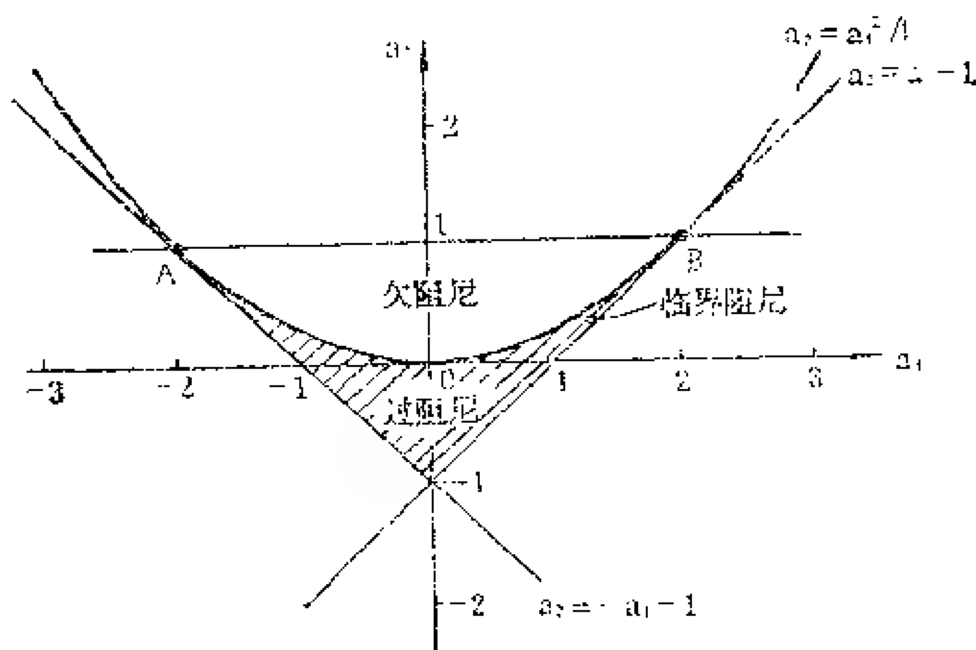


图2-7 AR(2)过程的特性与系数分布图

现在来看$AR(2)$过程的方差和自相关函数，在满足渐近平稳的条件下，在式(2-3-25)两边乘以x_{t-r}并取期望可得

$$E[x_t x_{t-r}]+a_1E[x_{t-1}x_{t-r}]+a_2E[x_{t-2}x_{t-r}]=E[\varepsilon_t x_{t-r}]$$
$$r=0,1,2,\cdots \tag{2-3-34}$$

由于$r\geqslant0$时$x_{t-r}=\varepsilon_{t-r}+(\varepsilon_{t-s};\ s>r$的线性函数)(见式(2-3-29),因此$E(x_{t-r}\varepsilon_t)=0(r\geqslant1)$,且$E(x_t\varepsilon_t)=E(\varepsilon_t^2)=\sigma_\varepsilon^2$,故由式(2-3-34)不难得出

$$r=0\quad \sigma_x^2[1+a_1\rho(1)+a_2\rho(2)]=\sigma_\varepsilon^2 \tag{2-3-35a}$$

$$r=1\quad \rho(1)+a_1\rho(0)+a_2\rho(1)=0 \tag{2-3-35b}$$

$$r=2\quad \rho(2)+a_1\rho(1)+a_2\rho(2)=0 \tag{2-3-35c}$$

$\cdots$

$$r>2\quad \rho(r)+a_1\rho(r-1)+a_2\rho(r-2)=0 \tag{2-3-35d}$$

解式(2-3-35)b,c得

$$\rho(1)=-a_1/(1+a_2) \tag{2-3-36}$$

$$\rho(2)=-a_2+a_1^2/(1+a_2) \tag{2-3-37}$$

代入式(2-3-35)a得

$$\sigma_x^2=\frac{(1+a_2)}{(1-a_2)(1+a_1+a_2)(1-a_1+a_2)}\sigma_\varepsilon^2 \tag{2-3-38}$$

在式(2-3-33)条件满足的情况下,σ_x^2显然总是正的。

由式(2-3-35)d还可看出$r>2$时的ρ值可在$\rho(1)$和$\rho(2)$的基础上递推计算,也可用解二阶齐次差分方程的方法写出

$$\rho(r)=B_1\lambda_1^r+B_2\lambda_2^r\qquad r\geqslant0 \tag{2-3-39}$$

然后根据边界条件确定系数B_1,B_2、即由

$$\rho(0)=B_1+B_2=1$$

$$\rho(1)=B_1\lambda_1+B_2\lambda_2=-a_1/(1+a_2)$$

解出B_1,B_2再代入式(2-3-39)得

$$\rho(r)=\frac{(1-\lambda_2^2)\lambda_1^{r+1}-(1-\lambda_1^2)\lambda_2^{r+1}}{(\lambda_1-\lambda_2)(1+\lambda_1\lambda_2)}\qquad r\geqslant0 \tag{2-3-40}$$

图2-8给出四个系数分布不同的$AR(2)$模型的$\rho(r)$理论值的图形。
(1) $x_t-0.5x_{t-1}-0.3x_{t-2}=\varepsilon_t$ (2) $x_t+0.5x_{t-1}-0.3x_{t-2}=\varepsilon_t$
(3) $x_t+x_{t-1}+0.6x_{t-2}=\varepsilon_t$ (4) $x_t-x_{t-1}+0.6x_{t-2}=\varepsilon_t$

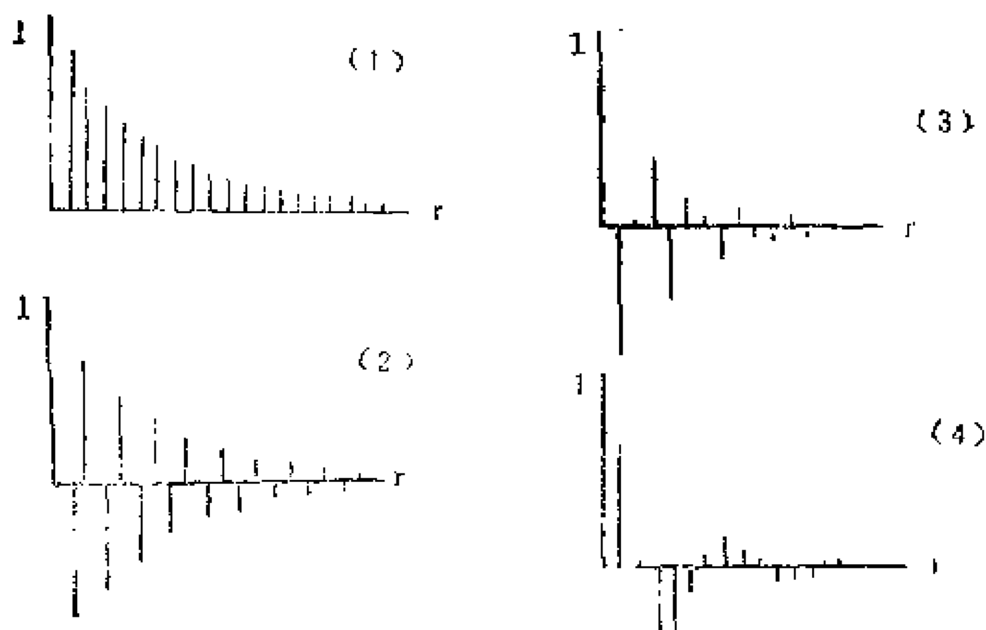


图2-8　AR (2) 过程的自相关

其中(1)和(2)的根属于实根($a_2 \leqslant a_1^2/4$)，(1)中以正实根为主($\lambda_1=0.85$, $\lambda_2=-0.35$)，ρ呈单调指数衰减，(2)中以负实根为主($\lambda_1=0.35$, $\lambda_2=-0.85$)，ρ正负相间地指数衰减，(3)和(4)均属复根($a_2>a_1^2/4$)，(3)中$\rho(1)$和$\rho(0)$必反号，而(4)中开始有2个ρ值均为正，复根情况的特点是$\rho(r)$振荡式衰减，即具有"伪周期"如果把根写成$\lambda_1=\sqrt{a_2}\,e^{j\omega}$和$\lambda_2=\sqrt{a_2}\,e^{-j\omega}$，其中

$$\cos\omega=-a_1/(2\sqrt{a_2}) \tag{2-3-41}$$

则代入式(2-3-40)经化简可得$\rho(r)$的另一种写法

$$\rho(r)=a_2^{r/2}\left[\frac{\sin(r\omega+\psi)}{\sin\psi}\right] \tag{2-3-42}$$

其中　　$\tan\psi=[(1+a_2)/(1-a_2)]\tan\omega$

由式(2-3-42)可见，尽管$\rho(r)$呈指数衰减(因$|a_2|<1$，而$a_2^{r/2}$相当于衰减振幅)，但[·]项是振荡型的，其周期为$2\pi/\omega$，此即ρ的"伪周期"，当$a_2\to1$时$\psi\to\pi/2$，从而$\rho(r)=\cos r\omega$，这时ρ具

有真正的周期，它表示随机过程本身具有这种周期。ρ的伪周期也表示x_t中具有某种伪周期现象——相距为一两个“周期”时，过程的值也较“相似”，但相隔好几个“周期”之后，彼此就不像了。随机性的冲击作用加在弱阻尼的系统中常产生这种伪周期过程，$2\pi/\omega$是这种系统的自然频率，ε_t相当于一系列的随机冲击。

4. n阶自回归过程 称$\{x_t\}$为n阶自回归过程$AR(n)$，如果它满足差分方程

$$x_t+a_1x_{t-1}+\cdots+a_nx_{t-n}=\varepsilon_t \tag{2-3-43}$$

其中a_1，a_2，…，a_n为常数，$\{\varepsilon_t\}$为纯随机过程，上式亦可写成

$$\Phi(B)x_t=\varepsilon_t \tag{2-3-44}$$

其中 $$\Phi(B)=1+a_1B+a_2B^2+\cdots+a_nB^n \tag{2-3-45}$$

它的解由两部分组成，即

$$x_t=g(t)+\Phi^{-1}(B)\varepsilon_t \tag{2-3-46}$$

其中$g(t)$是齐次方程$\Phi(B)x_t=0$的解，即

$$g(t)=A_1\lambda_1^t+A_2\lambda_2^t+\cdots+A_n\lambda_n^t \tag{2-3-47}$$

其中λ_1，λ_2，…，λ_n是$f(Z)=Z^n+a_1Z^{n-1}+\cdots+a_{n-1}Z+a_n$的根（设诸根不等），渐近平稳要求$t\to\infty$时$g(t)\to0$，因此$f(Z)$所有的根都在$Z$平面单位圆内，即$|\lambda_i|<1$，$i=1$，2，…，$n$，或者说$\Phi(B)$的根都在$B$平面的单位圆外。

若将$\Phi(B)$按因式分解，即$\Phi(B)=\prod\limits_{i=1}^{n}(1-\lambda_iB)$，则式(2-3-46)中第二项可写成

$$\Phi^{-1}(B)\varepsilon_t=\left[\sum_{i=0}^{n}\frac{c_i}{1-\lambda_iB}\right]\varepsilon_t \tag{2-3-48}$$

其中c_i为常数，当$|\lambda_i|<1(i=1, 2, \cdots, n)$的情况下，上式的每

一项都可展成一个（均方）收敛级数，因此式(2-3-48)代表x_t的平稳解。

当平稳条件满足，且$E[\varepsilon_t]=0$(从而$E[x_t]=0$)，在式(2-3-43)两边乘以$x_{t-r}(r\geqslant 0)$并取期望可得

$$r=0 \quad \sigma_x^2[1+a_1\rho(1)+a_2\rho(2)+\cdots+a_n\rho(n)]=\sigma_\varepsilon^2 \quad (2-3-49)$$

$$r=1 \quad \rho(1)+a_1\rho(0)+a_2\rho(1)+\cdots+a_n\rho(n-1)=0 \quad (2-3-50)$$

$$\cdots$$

$$r\geqslant n \quad \rho(r)+a_1\rho(r-1)+a_2\rho(r-2)+\cdots+a_n\rho(r-n)=0 \quad (2-3-51)$$

如果考虑到ρ是r的偶函数，则自式(2-3-50)起方程式中ρ的自变量均按递减顺序排列，是$\rho(r)$的齐次差分方程，它们称作为耶尔—瓦克尔(*Yule-Walker*)方程。当$f(Z)$中无重根时

$$\rho(r)=B_1\lambda_1^{|r|}+B_2\lambda_2^{|r|}+\cdots+B_n\lambda_n^{|r|} \quad (2-3-52)$$

其中B_1，B_2，…，B_n取决于边界条件$\rho(0)$，$\rho(1)$，…，$\rho(n-1)$。若λ_i全为实根，则$\rho(r)$指数衰减，快慢取决于主根，若λ_i中有复根，则$\rho(r)$振荡衰减。

关于$AR(n)$过程是否满足平稳条件，比较简便的方法是利用朱利(*Jury*)准则（程序见附录二JURYC）：实系数多项式

$$f(Z)=a_0Z^n+a_1Z^{n-1}+\cdots+a_{n-1}Z+a_n \quad a_0>0 \quad (2-3-53)$$

的根全部在Z平面单位圆以内的充分必要条件。为：

$$(1)f(1)>0 \quad (2-3-54)$$

$$(2)f(-1)\begin{cases}<0 & n\text{奇数}\\ >0 & n\text{偶数}\end{cases} \quad (2-3-55)$$

$$(3)|\alpha_i|<1 \quad i=0,\ 1,\ 2,\ \cdots,\ n-2 \quad (2-3-56)$$

其中$\alpha_0=a_n/a_0$，$\alpha_1=b_1/b_n$，$\alpha_2=c_2/c_n$，$\alpha_3=d_3/d_n$，…，

$\alpha_{n-2}=q_{n-2}/q_n$。

b_k，c_k，d_k，…，q_k是按照下述方法求得的

$$b_k=\begin{vmatrix} a_n & a_{n-k} \\ a_0 & a_k \end{vmatrix} \qquad k=1,\ 2,\ \cdots,\ n$$

$$c_k=\begin{vmatrix} b_0 & b_{n-k-1} \\ b_1 & b_k \end{vmatrix} \qquad k=2,\ 3,\ \cdots,\ n$$

$$d_k=\begin{vmatrix} c_n & c_{n-k-2} \\ c_2 & c_k \end{vmatrix} \qquad k=3,\ 4,\ \cdots,\ n$$

$\vdots$

直到只取$k=n-3$，$n-2$，$n-1$，n得相应的p_k，而

$$q_{n-2}=\begin{vmatrix} p_n & p_{n-1} \\ p_{n-3} & p_{n-2} \end{vmatrix},\ q_{n-1}=\begin{vmatrix} p_n & p_{n-2} \\ p_{n-3} & p_{n-1} \end{vmatrix},\ q_n=\begin{vmatrix} p_n & p_{n-3} \\ p_{n-3} & p_n \end{vmatrix}$$

以上诸条件只要有一项不成立则必有根在单位圆上或圆外。由于(1)(2)条件简单，一般先检查，如果有不满足的情况则(3)就无需检查了。

例：$f(Z)=10Z^3-41Z^2+54Z-5=0$，可求得$f(1)=18>0$，$f(-1)=-110<0$，$|\alpha_0|=0.5<1$，$|\alpha_1|=|335/75|>1$，不满足。

5. **滑动平均过程** 另一种常见的过程是滑动平均过程。

$\{x_t\}$ 称为m阶滑动平均过程MA(m)，如果它可表为

$$x_t=b_0\varepsilon_t+b_1\varepsilon_{t-1}+\cdots+b_m\varepsilon_{t-m} \tag{2-3-57}$$

其中b_0，b_1，…为常数，$\{\varepsilon_t\}$ 为纯随机过程，上式亦可写成

$$x_t=\Theta(B)\varepsilon_t \tag{2-3-58}$$

其中
$$\Theta(B)=b_0+b_1B+\cdots+b_mB^m \tag{2-3-59}$$

不失一般性，可取$b_0=1$，因为如果$b_0\neq 1$，可定义新的过程$\varepsilon_t'=b_0\varepsilon_t$，它也是纯随机过程，这样用$\varepsilon_t'$来表示$x_t$时，相应差分方程右边第一项系数为1。另外，也常假定ε_t的方差为1，因若$\sigma_\varepsilon^2\neq 1$，则可定义$\varepsilon_t''=\varepsilon_t/\sigma_\varepsilon$，而$\varepsilon_t''$仍为纯随机过程，不过我们无法同时做上述两种变换来达到两个目的。

MA过程和AR过程不同，它是 $\{\varepsilon_t\}$ 过程的现在和有限范围内的过去值的线性组合，所以ε_t只影响x_t的m个未来值。换句话

说，相隔时间大于m的x_t中所包含的ε_t已全部更新，所以相隔时间超过m的两个x_t是不相关的，因此MA过程的$\rho(r)$在m点之后将"截尾"——$\rho(r)=0$，$|r|>m$。由于$\{x_t\}$是无关且平稳的$\{\varepsilon_t\}$的线性组合，显然$\{x_t\}$总是平稳的而不论b_i的值如何。不难得出$\{x_t\}$的方差为

$$\sigma_x^2=\sigma_\varepsilon^2(b_0^2+b_1^2+\cdots+b_m^2) \tag{2-3-60}$$

而若$E[\varepsilon_t]=0$，则$E[x_t]=0$，故有

$$\rho(r)=E[x_t x_{t-r}]/\sigma_x^2$$

$$=\begin{cases}\sigma_\varepsilon^2(b_r b_0+b_{r+1}b_1+\cdots+b_{m-r})/\sigma_x^2 & 0\leqslant r\leqslant m\\ 0 & r>m\end{cases} \tag{2-3-61}$$

或

$$\rho(r)=\begin{cases}(\sum\limits_{i=r}^{m} b_i b_{i-r})/(\sum\limits_{i=0}^{m} b_i^2) & 0\leqslant r\leqslant m\\ 0 & r>m\\ \rho(-r) & r<0\end{cases} \tag{2-3-62}$$

图2-9给出一些MA(1)和MA(2)过程在各种不同系数情况下的自相关函数的图例。其中

(1)a) $x_t=\varepsilon_t+0.8\varepsilon_{t-1}$　　(1)b) $x_t=\varepsilon_t-0.8\varepsilon_{t-1}$

(2)a) $x_t=\varepsilon_t+0.5\varepsilon_{t-1}-0.3\varepsilon_{t-2}$　　(2)b) $x_t=\varepsilon_t-0.5\varepsilon_{t-1}-0.3\varepsilon_{t-2}$

(2) c) $x_t=\varepsilon_t+\varepsilon_{t-1}+0.6\varepsilon_{t-2}$　(2) d) $x_t=\varepsilon_t-\varepsilon_{t-1}+0.6\varepsilon_{t-2}$

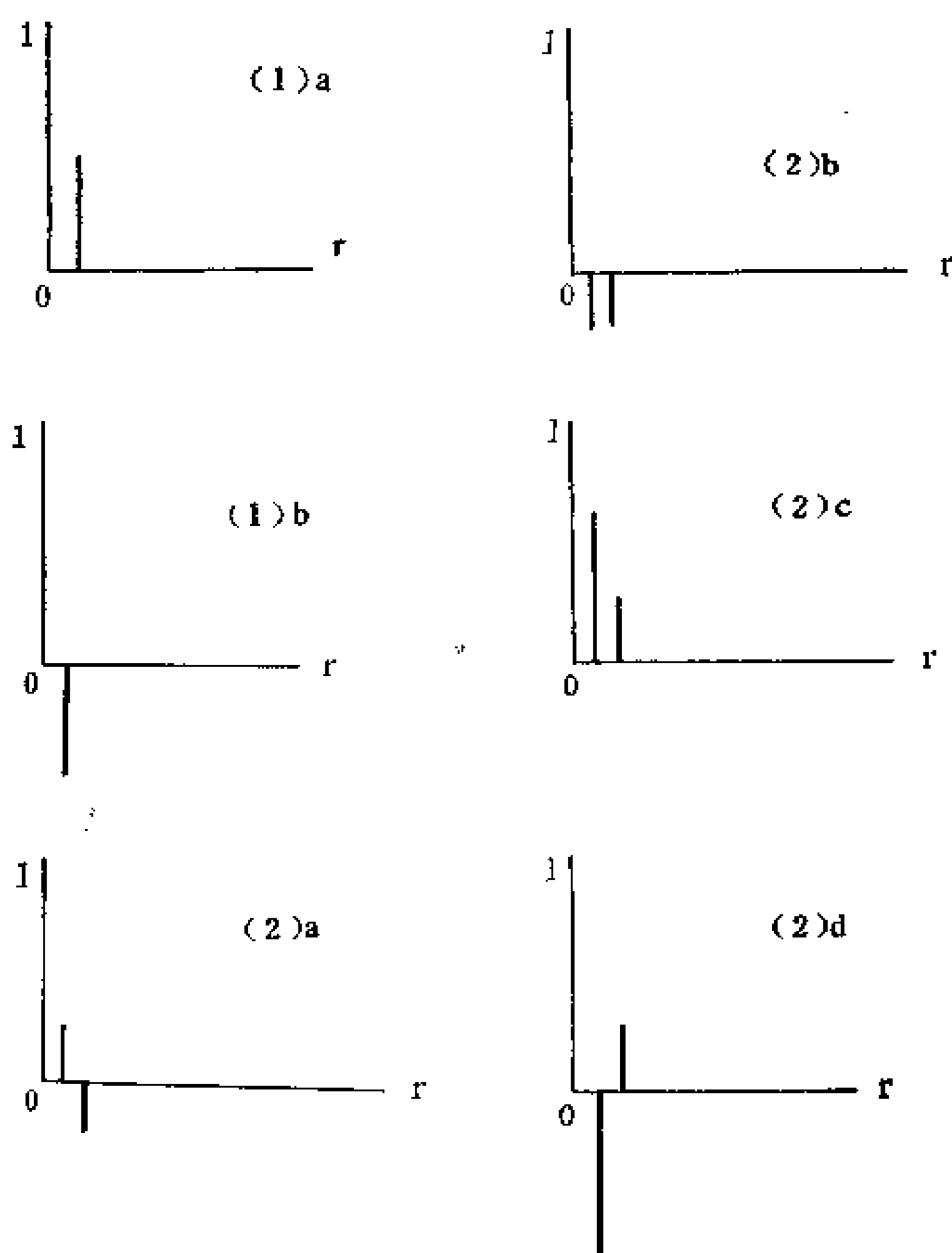


图2-9 滑动平均过程的自相关

6. 自回归滑动平均过程,格林函数和逆函数 更一般的过程是用(n阶)自回归(m阶)滑动平均的混合模型来描述,称为ARMA(n,m)过程,它满足

$$x_t+a_1x_{t-1}+\cdots+a_nx_{t-n}=b_0\varepsilon_t+b_1\varepsilon_{t-1}+\cdots+b_m\varepsilon_{t-m} \tag{2-3-63}$$

其中a_i，b_i为常数，$\{\varepsilon_t\}$为纯随机过程，不失一般性可取$b_0=1$。AR和MA过程显然都是$ARMA$的特殊情况。图2-10是ARMA(2.2) $x_t+1.45x_{t-1}+0.6x_{t-2}=\varepsilon_t-0.2\varepsilon_{t-1}-0.1\varepsilon_{t-2}$（$\varepsilon_t$是零均值单位方差的正态随机过程）的200点观察值曲线。

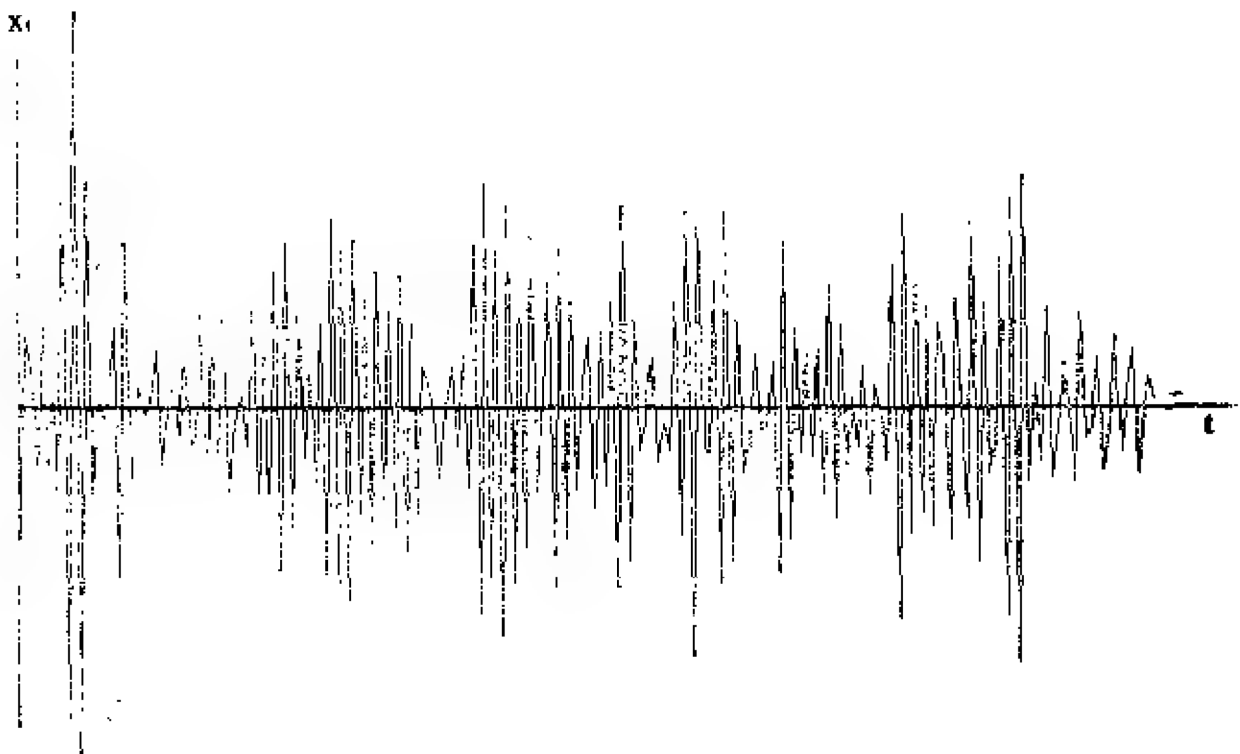


图2-10　一个ARMA（2,2）过程的实现

ARMA(n,m)的平稳性条件和$AR(n)$相同，利用后移算子可将(2-3-62)式可写成

$$\Phi(B)x_t=\Theta(B)\varepsilon_t \tag{2-3-63}a$$

如果AR部分的系数满足平稳性条件，即$\Phi(B)=1+a_1B+\cdots+a_nB^n$的根全在$B$平面单位圆之外，则$x_t$的平稳解为

$$x_t=\Phi^{-1}(B)\Theta(B)\varepsilon_t \tag{2-3-64}$$

将$\Phi^{-1}(B)$展成B的级数（在平稳条件满足时，这种展开是存在的），则上式右边的算子式为

$$G(B)=\Phi^{-1}(B)\Theta(B)=\sum_{i=0}^{\infty}G_iB^i \tag{2-3-65}$$

其中系数序列G_0，$G_1, G_2, \cdots$称为格林(Green)函数。例如ARMA(2,1)过程：$x_t - 1.3x_{t-1} + 0.4x_{t-2} = \varepsilon_t - 0.4\varepsilon_{t-1}$有

$$G(B) = \frac{1-0.4B}{1-1.3B+0.4B^2} = 1+0.9B+0.77B^2+0.64B^3+\cdots$$

可见格林函数为 $\{G_i\} = \{G_0, G_1, G_2, \cdots,\} = \{1, 0.9, 0.77, 0.64, \cdots\}$

利用式(2-3-64)可将x_t用 $\{\varepsilon_t\}$ 的现在和过去的值（一般为无限多项）来表达，即

$$x_t = \left(\sum_{i=0}^{\infty} G_i B^i\right)\varepsilon_t = G_0\varepsilon_t + G_1\varepsilon_{t-1} + G_2\varepsilon_{t-2} + \cdots$$

$$= \sum_{i=0}^{\infty} G_i \varepsilon_{t-i} \tag{2-3-66}$$

显然它可以看作是无限阶的MA过程。平稳过程x_t的这种表达形式称为它的"传递形式"，作为传递形式的系数——格林函数实际上是权重函数，它表示i个时间单位以前的ε_t对现在x_t值影响的程度。

对于AR和$ARMA$过程而言，格林函数 $\{G_i\}$ 有无限多项，但对于$MA(m)$过程来说，$G_0 = b_0$，$G_1 = b_1$，…，$G_m = b_m$，而当$i > m$时$G_i = 0$。

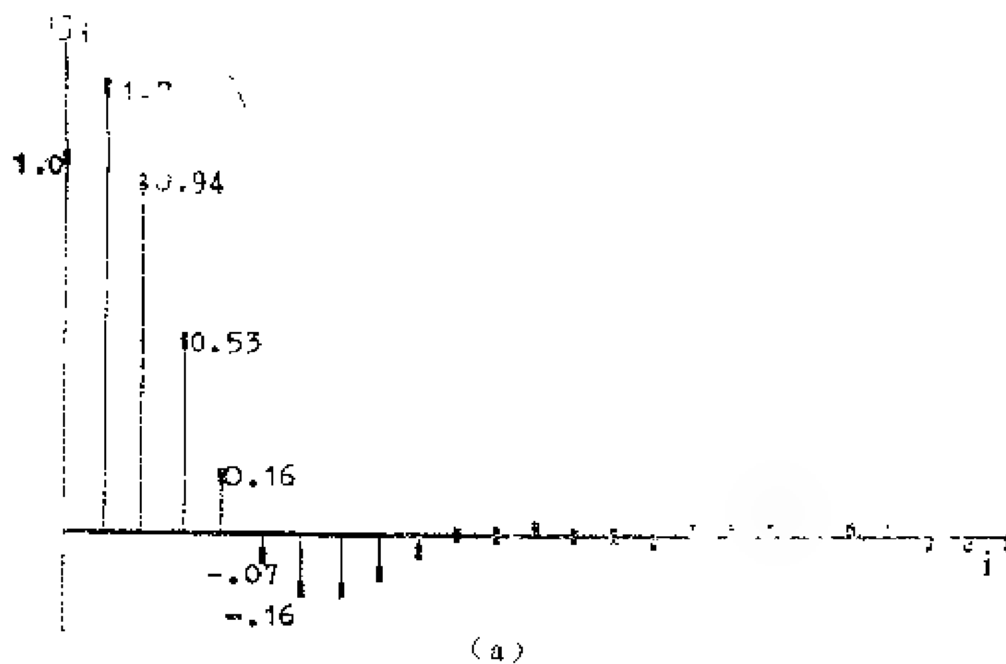
通常取滑动平均的第一项系数$b_0 = 1$，即

$$x_t + a_1x_{t-1} + \cdots + a_nx_{t-n} = \varepsilon_t + b_1\varepsilon_{t-1} + \cdots + b_m\varepsilon_{t-m} \tag{2-3-67}$$

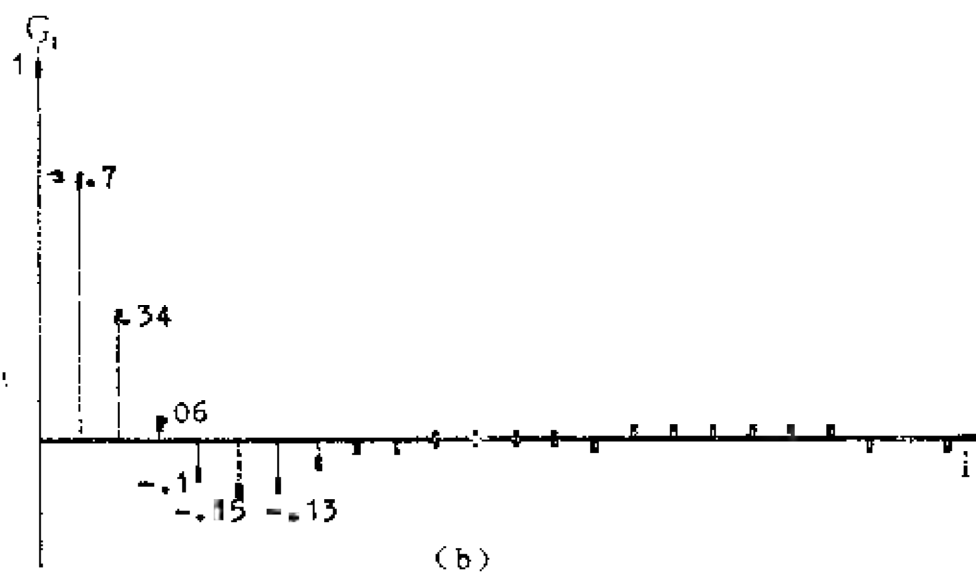
这时$G_0 \equiv 1$，而将$\Phi(B)G(B) = \Theta(B)$两边展成多项式后，比较B同幂项的系数可得G_i的递推公式

$$G_i = b'_i - \sum_{j=1}^{i} a'_j G_{i-j} \qquad G_0 \equiv 1 \tag{2-3-68}$$

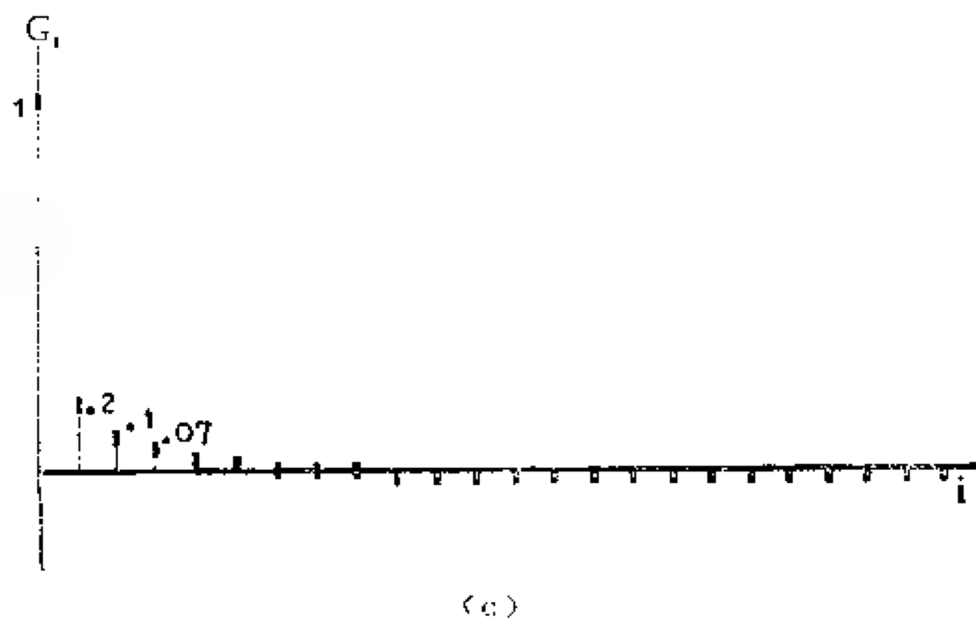
其中 $$a'_j = \begin{cases} a_j & 1 \leqslant j \leqslant n \\ 0 & j > n \end{cases} \quad b'_j = \begin{cases} b_j & 1 \leqslant j \leqslant m \\ 0 & j > m \end{cases} \tag{2-3-69}$$



(a)



(b)



(c)

图2-11　格林函数示例

图2-11画出一些给定过程的格林函数图形，这些过程都是平稳的，因此相应的格林函数都是衰减的（$i\to\infty$时$G_i\to 0$）。其中$a)a_1=-1.2$，$a_2=0.5$($AR(2)$过程)；$b)a_1=-1.2$，$a_2=0.5$，$b_1=-0.5$；$c)a_1=-0.7$，$a_2=0.03$，$b_1=-0.5$

下面利用式(2-3-66)导出过程相关函数的关系式。先在式(2-3-63)两边乘以x_{t-r}并取期望，

$$E[x_{t-r}x_t]+a_1E[x_{t-r}x_{t-1}]+\cdots+a_nE[x_{t-r}x_{t-n}]$$
$$=E[x_{t-r}\varepsilon_t]+b_1E[x_{t-r}\varepsilon_{t-1}]+\cdots+b_mE[x_{t-r}\varepsilon_{t-m}]$$

定义ε_t和x_t的互相关为$R_{x\varepsilon}(k)=E[x_t\varepsilon_{t+k}]$，则上式可写成

$$R_x(r)+a_1R_x(r-1)+\cdots+a_nR_x(r-n)$$
$$=R_{x\varepsilon}(r)+b_1R_{x\varepsilon}(r-1)+\cdots+b_mR_{x\varepsilon}(r-m) \qquad (2\text{-}3\text{-}70)$$

利用式(2-3-66)有

$$R_{x\varepsilon}(k)=E\left[\sum_{i=0}^{\infty}G_i\varepsilon_{t-i}\varepsilon_{t+k}\right]=\sum_{i=0}^{\infty}G_iE[\varepsilon_{t-i}\varepsilon_{t+k}]$$
$$=\begin{cases}G_{-k}\sigma_\varepsilon^2 & k\leqslant 0\\ 0 & k>0\end{cases} \qquad (2\text{-}3\text{-}71)$$

将此式代入式(2-3-70)右边中的$R_{x\varepsilon}$，并考虑到ρ为偶函数，可得

$$r=0\quad \sigma_x^2[1+a_1\rho(1)+a_2\rho(2)+\cdots+a_n\rho(n)]=(1+b_1G_1+\cdots+b_mG_m)\sigma_\varepsilon^2$$
$$r=1\quad \sigma_x^2[\rho(1)+a_1\rho(0)+a_2\rho(1)+\cdots+a_n\rho(n-1)]=(b_1+b_2G_1+\cdots+b_mG_{m-1})\sigma_\varepsilon^2$$
$$\vdots \qquad (2\text{-}3\text{-}72)$$
$$r=m\quad \sigma_x^2[\rho(m)+a_1\rho(m-1)+\cdots+a_n\rho(n-m)]=b_m\sigma_\varepsilon^2$$
$$r>m\quad \rho(r)+a_1\rho(r-1)+\cdots+a_n\rho(n-r)=0$$

当$n\leqslant m+1$且$r\geqslant m+1$，或当$n>m+1$且$r\geqslant n$时$\rho(r)$均满足齐次差分方程关系，式中ρ的变量为正且递减，换言之，当$r\geqslant\max(n,$

$m+1$)时$\rho(r)$和纯AR模型中x_t的齐次差分方程关系式相同，由它可以解出$\rho(r)$。实际计算可先按式(2-3-68)求出$G_1, \cdots, G_m$，代入式(2-3-72)，在解出$\rho(1)$，…，$\rho(n)$之后用递推方法计算其余的ρ值，详见附录二中的程序$AUTOCR$及其说明。

例1. ARMA(1,1)过程$x_t + a x_{t-1} = \varepsilon_t + b\varepsilon_{t-1}$，$\varepsilon_t$均值为零，方差为$\sigma_\varepsilon^2$，若$|a|<1$则有

$$x_t = G(B)\varepsilon_t$$

其中 $$G(B) = \frac{1+bB}{1+aB} = G_0 + G_1 B + G_2 B^2 + \cdots$$

$$G_0 = 1,\ G_1 = b-a,\ G_2 = -a(b-a),\ \cdots$$

根据式(2-3-72)有

$$\sigma_x^2[1+a\rho(1)] = (1+b^2-ab)\sigma_\varepsilon^2 \qquad r=0$$

$$\sigma_x^2[\rho(1)+a] = b\sigma_\varepsilon^2 \qquad r=1$$

$$\rho(r)+a\rho(r-1)=0 \qquad r=2,\ 3,\ 4,\ \cdots$$

可解得 $$\sigma_x^2 = \sigma_\varepsilon^2(1+b^2-2ab)/(1-a^2)$$

$$\rho(1) = [b(1+a^2)-a(1+b^2)]/(1+b^2-2ab)$$

$\rho(2)$，$\rho(3)$，…可以递推或解齐次差分方程求得。

例2. ARMA(2,1)过程$x_t + a_1 x_{t-1} + a_2 x_{t-2} = \varepsilon_t + b\varepsilon_{t-1}$，可求得

$$G(B) = \frac{1+bB}{1+a_1 B + a_2 B^2}$$

$$= 1 + (b-a_1)B + [a_1(a_1-b) - a_2]B^2 + \cdots$$

故 $G_0 = 1$，$G_1 = b-a_1$，$G_2 = a_1(a_1-b)-a_2$，…

由式(2-3-72)有

$$\sigma_x^2[1+a_1\rho(1)+a_2\rho(2)] = (1+bG_1)\sigma_\varepsilon^2 = (1+b^2-a_1 b)\sigma_\varepsilon^2 \qquad r=0$$

$$\sigma_x^2[\rho(1)+a_1+a_2\rho(1)] = b\sigma_\varepsilon^2 \qquad r=1$$

$$\rho(r)+a_1\rho(r-1)+a_2\rho(r-2)=0 \qquad r=2,3,4\cdots$$

将$\rho(2)=-[a_1\rho(1)+a_2]$代入第一式有

$$\sigma_x^2[1-a_2+a_1(1-a_2)\rho(1)]=(1+b^2-a_1b)\sigma_\varepsilon^2$$

由此和第二式可以解得

$$\sigma_x^2=\sigma_\varepsilon^2\frac{\begin{vmatrix}1-a_1b+b^2 & a_1(1-a_2)\\ b & 1+a_2\end{vmatrix}}{\begin{vmatrix}1-a_2^2 & a_1(1-a_2)\\ a_1 & 1+a_2\end{vmatrix}}$$

$$\rho(1)=\frac{\begin{vmatrix}1-a_2^2 & 1-a_1b+b^2\\ a_1 & b\end{vmatrix}}{\begin{vmatrix}1-a_1b+b^2 & a_1(1-a_2)\\ b & 1+a_2\end{vmatrix}}$$

$\rho(2)$，$\rho(3)$，…可按递推或解齐次差分方程求得。

自相关函数也可以用格林函数直接表达，这时利用ARMA过程的传递形式有

$$R_x(r)=E[x_t x_{t+r}]=E[\sum_{i=0}^{\infty}G_i\varepsilon_{t-i}\sum_{l=0}^{\infty}G_l\varepsilon_{t+r-l}]$$

$$=\sum_{i=0}^{\infty}G_i\sum_{l=0}^{\infty}G_l R_\varepsilon(r+i-l) \tag{2-3-73}$$

考虑到R_ε是方差为σ_ε^2的白噪声ε_t的自相关，上式在$l=r+i$时$R_\varepsilon(0)=\sigma_\varepsilon^2$，其他$l$处$R_\varepsilon$均为零，故得

$$R_x(r)=\sigma_\varepsilon^2\sum_{i=0}^{\infty}G_iG_{i+r} \tag{2-3-74}$$

$$\rho_x(r)=(\sigma_\varepsilon^2/\sigma_x^2)\sum_{i=0}^{\infty}G_iG_{i+r} \tag{2-3-75}$$

除了传递形式外，x_t的又一种表达形式是"逆转形式"，这是和式(2-3-64)的做法相反，将x_t和ε_t的关系式写成

$$\Theta^{-1}(B)\Phi(B)x_t=\varepsilon_t \qquad (2\text{-}3\text{-}76)$$

这里假定$\Theta(B)=1+b_1B+\cdots+b_mB^m$的根全部处在$B$平面单位圆之外，因此$\Theta^{-1}(B)$展成的$B$的级数存在，令式(2-3-76)左边的算子式为$I(B)$，则

$$I(B)=\Theta^{-1}(B)\Phi(B)=1-\sum_{i=1}^{\infty}I_iB^i \qquad (2\text{-}3\text{-}77)$$

因此把白噪声ε_t表达成以x_t为项的级数

$$\varepsilon_t=I(B)x_t=x_t-I_1x_{t-1}-I_2x_{t-2}-\cdots=\sum_{j=0}^{\infty}(-I_j)B^jx_t$$

$$(2\text{-}3\text{-}78)$$

称上式系数序列$\{I_i\}$ $(I_0\equiv-1)$为逆函数，显然，对于AR(n)过程而言，其逆函数只到I_n为止，即$I_1=-a_1$，$I_2=-a_2$，…，$I_n=-a_n$，$I_i=0(i>n)$。而MA和一般的ARMA过程的逆函数有无穷多项，过程是否能写成逆转形式，或者说逆函数是否存在，取决于滑动平均部分的参数b_j，如同判别$\Phi(B)$是否符合平稳性条件一样，可以用朱利准则判别$\Theta(B)$是否具备可逆性条件。

图2-12给出$x_t-1.2x_{t-1}+0.5x_{t-2}=\varepsilon_t-0.5\varepsilon_{t-1}$过程的逆函数图形。

由$\Phi(B)=\Theta(B)I(B)$可以导出I_i的递推算式如下：

$$I_i=-a'_i-\sum_{j=1}^{i}b'_jI_{i-j} \qquad I_0\equiv-1, \qquad (2\text{-}3\text{-}79)$$

其中a'_i和b'_j的含义同式(2-3-69)。

以上讨论可以看出，当过程为平稳且可逆时各种模型有如下转换关系：

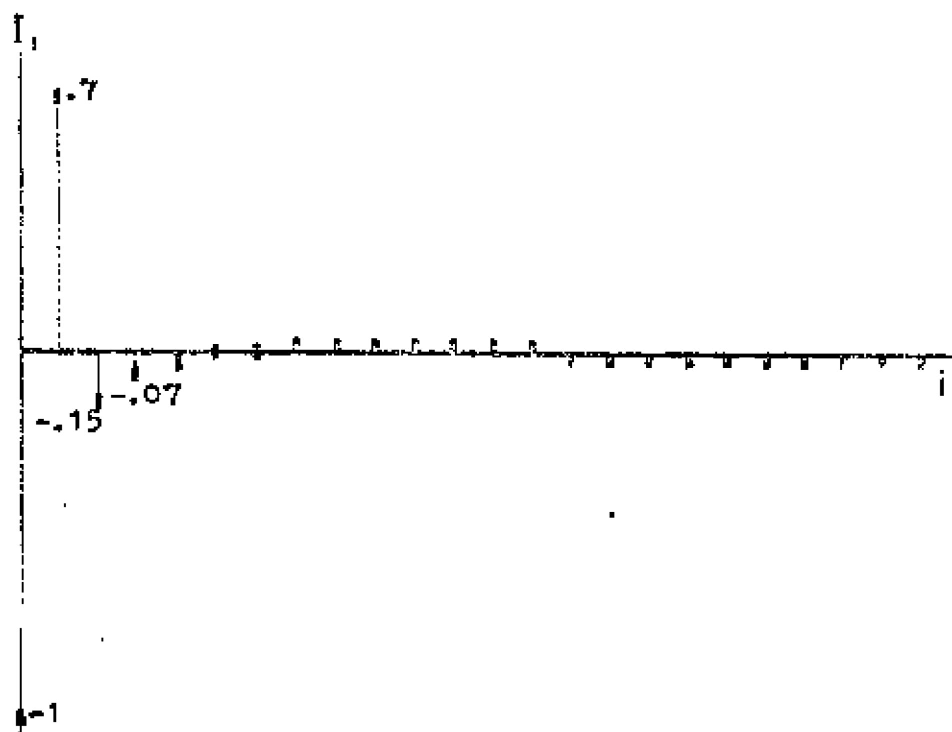


图2-12　逆函数示例

AR(有限阶)⇒MA(无限阶)

MA(有限阶)⇒AR(无限阶)

ARMA(有限阶)⇒AR(无限阶)

ARMA(有限阶)⇒MA(无限阶)

ARMA模型足以代表相当大范围的各种平稳随机过程，即适当选择有限个参数($n+m$个)便可以对序列作较好的拟合，这时用纯AR或纯MA模型来描述则需用很多个参数才能达到较好的拟合，换言之，ARMA所用的参数比较"节约"。不过从拟合计算的复杂性来说，*AR*模型比较简便，所以在阶数不是很高，或要求计算快速简单的情况下，还是愿意采用*AR*模型。

2.4　平稳随机过程的频率域表示

前面用差分方程或相关函数等描述一个平稳随机过程是属于

"时间域"的方法，在频率域内描述平稳随机过程是用不同频率的正弦和余弦函数的叠加（或积分），这时要利用富氏变换的概念，我们假定读者对这种变换已初步了解，在附录一中，还有关于富氏变换性质的归纳可资参考。

1. **频谱分析的基本概念** 我们知道确定性的周期函数在一定条件下可以看成由 $\{\cos ks,\ \sin ks,\ k=0,1,2\cdots\}$ 这样一些"元素"所构成，这是数学家富里叶(Fourier)首先提出的基本思想。一个周期为$2T_0$的信号$x(t)$(即$x(t+k2T_0)\equiv x(t)$)就是该信号的均值（直流分量）和频率为$1/2T_0$的基波($f_0=1/2T_0$称基波频率)及其谐波(频率为$k/2T_0$，$k=2,\ 3,\ 4\cdots$)的正弦函数和余弦函数的总和，这就是富氏级数展开：

$$x(t)=\frac{a_0}{2}+\sum_{k=1}^{\infty}\left(a_k\cos\frac{\pi kt}{T_0}+b_k\sin\frac{\pi kt}{T_0}\right) \tag{2-4-1}$$

其中

$$a_k=\frac{1}{T_0}\int_{-T_0}^{T_0}x(t)\cos\frac{\pi kt}{T_0}dt \qquad k=0,1,2,\cdots$$

$$b_k=\frac{1}{T_0}\int_{-T_0}^{T_0}x(t)\sin\frac{\pi kt}{T_0}dt \qquad k=1,2,\cdots \tag{2-4-2}$$

若将正余弦函数用指数形式表示，并令

$A_0=a_0/2$，$A_k=(a_k-jb_k)/2$，$A_{-k}=(a_k+jb_k)/2$(其中$j=\sqrt{-1}$，$k=1,\ 2,\ \cdots$)，则得复数形式的富氏级数：

$$x(t)=\sum_{k=-\infty}^{\infty}A_ke^{j\pi kt/T_0}=\sum_{k=-\infty}^{\infty}A_ke^{j2\pi f_kt} \tag{2-4-3}$$

其中$f_k=k/(2T_0)$，而富氏级数的系数A_k为

$$A_k=\frac{1}{2T_0}\int_{-T_0}^{T_0}x(t)e^{-j2\pi f_kt}dt \quad k=0,\pm1,\pm2,\cdots \tag{2-4-4}$$

在$x(t)$为实数的情况下有如下表达式

$$x(t)=A_0+2\sum_{k=1}^{\infty}|A_k|\cos(2\pi f_k t+\arg A_k) \tag{2-4-5}$$

其中

$$\begin{aligned}&|A_k|=\sqrt{a_k^2+b_k^2}/2\\&\arg A_k=\arctan(-b_k/a_k)\qquad k=1,2,\cdots\end{aligned} \tag{2-4-6}$$

对于确定性但非周期的$x(t)$，不能把它用适合于任何t的富氏级数表达，但可在有限区间(如$-T$到T)内使$x(t)$用一个富氏级数表示，而区间外则周期地重复上述区间内的规律。把这个构造的周期函数记作$x_{T\bar{p}}(t)$，即

$$\begin{aligned}&x_{T\bar{p}}(t)=x(t)\qquad -T\leqslant t\leqslant T\\&x_{T\bar{p}}(t+2nT)=x_{T\bar{p}}(t)\qquad n=1,2,\cdots\end{aligned} \tag{2-4-7}$$

在$-T\leqslant t\leqslant T$内

$$x(t)\equiv x_{T\bar{p}}(t)=\sum_{k=-\infty}^{\infty}A_k e^{j2\pi f_k t} \tag{2-4-8}$$

考虑到

$$\begin{aligned}A_k&=\frac{1}{2T}\int_{-T}^{T}x_{T\bar{p}}(t)e^{-j2\pi f_k t}dt\\&=\frac{1}{2T}\int_{-T}^{T}x(t)e^{-j2\pi f_k t}dt\end{aligned} \tag{2-4-9}$$

并令 $\frac{1}{2T}=f_k-f_{k-1}=\delta f_k$，有

$$x(t)=\sum_{k=-\infty}^{\infty}\left(\int_{-T}^{T}x(t)e^{-j2\pi f_k t}dt\right)e^{j2\pi f_k t}\,\delta f_k\qquad -T\leqslant t\leqslant T \tag{2-4-10}$$

当$T\to\infty$时$\delta f_k\to 0$，离散频率点（$\cdots, f_{-2}, f_{-1}, f_0, f_1, f_2, \cdots$）变成点的连续集合，式(2-4-10)右边变成积分，因此当$T\to\infty$时对所有的t有

$$x(t)=\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}x(t)e^{-j2\pi ft}dt\right)e^{j2\pi ft}df \qquad (2-4-11)$$

这就是确定性非周期函数的富氏积分。令

$$X(f)=\int_{-\infty}^{\infty}x(t)e^{-j2\pi ft}dt \qquad (2-4-12)$$

则

$$x(t)=\int_{-\infty}^{\infty}X(f)e^{j2\pi ft}df \qquad (2-4-13)$$

$X(f)$称为$x(t)$的富氏变换，记作$X(f)=\mathscr{F}[x(t)]$，而$x(t)=\mathscr{F}^{-1}[X(f)]$。当然上面是假定两个积分都存在，其条件是$x(t)$绝对可积——$\int_{-\infty}^{\infty}|x(t)|dt<\infty$.

将$x(t)$写成富氏积分形式依然表示它是sin和cos函数之和（取极限形式）。事实上，如果将$X(f)$写成$X(f)=g(f)-jh(f)$则式(2-4-13)可表达成

$$x(t)=\int_{-\infty}^{\infty}[g(f)\cos 2\pi ft+h(f)\sin 2\pi ft]df$$

这里关键是表达式(2-4-13)中的频率包含的是连续范围（$-\infty$，$+\infty$），这和富氏级数的离散频率不同。$X(f)$的物理意义可以解释为$x(t)$的总能量在连续频率域内的分布，因为总能量可以表达成下列积分

$$\int_{-\infty}^{\infty} x^2(t)dt=\int_{-\infty}^{\infty} x(t)\left[\int_{-\infty}^{\infty} X(f)e^{j2\pi ft}df\right]dt$$

$$=\int_{-\infty}^{\infty} X(f)\left[\int_{-\infty}^{\infty}x(t)e^{j2\pi ft}dt\right]df$$

$$=\int_{-\infty}^{\infty} X(f)X^*(f)df$$

$$=\int_{-\infty}^{\infty} |X(f)|^2df \tag{2-4-14}$$

所以$|X(f)|^2df$表示频率在f和$f+df$内$x(t)$的分量对总能量的贡献。$|X(f)|^2$表示某频率f处的能量密度。

以上对确定性函数的富氏分解做了启发性的说明。对于平稳随机过程来说却存在一些新问题：1)随机过程指的不仅是单个的函数，而是一个总体，即过程的无穷多个实现的集合。为此，在讨论时要从个别实现开始；2)对于一个实现，一般也不能用富氏级数表示，因为除非特殊情况，不能假定它是周期的；3)也不能用富氏积分来表示一个实现，因为不能保证它绝对可积。事实上平稳过程的实现几乎都不会在无限远处衰减到零。考虑上述情况，我们对某个实现加以“截取”以构造新函数：

$$x_T(t)=\begin{cases}x(t) & -T\leqslant t\leqslant T\\ 0 & \text{其他}\end{cases} \tag{2-4-15}$$

它和前面的x_{T_0}(式(2-4-7))不同，$x_T(t)$是非周期的，它在$(-T, T)$外为零而非周期重复，由于它绝对可积，可写成

$$x_T(t)=\int_{-\infty}^{\infty} X_T(f)e^{j2\pi ft}df \tag{2-4-16}$$

$$X_T(f)=\int_{-\infty}^{\infty} x_T(t)e^{-j2\pi ft}dt=\int_{-T}^{T} x(t)e^{-j2\pi ft}dt$$

(2-4-17)

$|X_T(f)|^2df$同样解释为f到$f+df$频率范围内的分量对$x_T(t)$总能量的贡献。

能否通过$T\to\infty$使$x_T(t)$在所有t处都和$x(t)$相同，从而用$\lim\limits_{T\to\infty}|X_T(f)|^2$描述$x(t)$的特性呢？不行，因为一般情况下平稳随机过程的富氏积分不存在，因此$\lim\limits_{T\to\infty}|X_T(f)|^2$不存在（非有限）。解决办法是引入功率来代替能量，因为$T\to\infty$时$|X_T(f)|^2$虽然趋于无穷，但功率收敛到有限值，即

$$\lim_{T\to\infty}|X_T(f)|^2/2T<\infty$$

而 $\lim\limits_{T\to\infty}\dfrac{|X_T(f)|^2}{2T}df=f$到$f+df$频率分量对$x(t)$总功率的贡献。

因此 $\lim\limits_{T\to\infty}\dfrac{|X_T(f)|^2}{2T}$——表示功率密度。

以上是对平稳随机过程的一个实现来讨论的，即$|X_T(f)|^2/2T$的值依赖于某一特定的具体实现。现在引入数学期望，令

$$G(f)=\lim_{T\to\infty}\left\{E\left[\frac{|X_T(f)|^2}{2T}\right]\right\} \tag{2-4-18}$$

则$G(f)df=f$到$f+df$间$x(t)$的分量对总功率的贡献的平均（对所有实现作平均）。

称$G(f)$为功率谱密度，它完全由$x(t)$的概率性质决定。但$G(f)$的求取并不一定要从它的原始定义出发，后面将证明$G(f)$是$\{x(t)\}$自协方差函数的富氏变换。如果知道$x(t)$为某种

$ARMA$过程，可以很方便地得出功率谱密度$G(f)$。

2．连续过程的功率谱和自协方差函数的关系 自协方差函数和功率谱特性之间为什么存在富氏变换的关系呢？由式(2-4-18)知某一频率处的功率取决于$|X_T(f)|^2$，可以证明$X_T(f)$和$x_T(t)$这一富氏变换对之间有如下关系

$$|X_T(f)|^2=\mathscr{F}\left[\int_{-\infty}^{\infty} x_T(u)x_T(u-\tau)du\right] \tag{2-4-19}$$

该式右边卷积的表达式和$E[x(t)x(t+\tau)]$很相像，所以如果从"功率"出发，必然会出现自协方差函数。下面先证明式(2-4-19)，令

$$k(\tau)=\int_{-\infty}^{\infty} x_T(u)x_T(u-\tau)du$$

则式(2-4-19)右边为

$$\mathscr{F}[k(\tau)]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-j2\pi f\tau}x_T(u)x(u-\tau)dud\tau$$

$$=\int_{-\infty}^{\infty} e^{-j2\pi fu}x_T(u)\left[\int_{-\infty}^{\infty} e^{j2\pi f(u-\tau)}x_T(u-\tau)d\tau\right]du$$

$$=X_T^*(f)\left[\int_{-\infty}^{\infty} e^{-j2\pi fu}x_T(u)du\right]=X_T^*(f)X_T(f)$$

$$=|X_T(f)|^2$$

式(2-4-19)得证，从而有

$$|X_T(f)|^2=\int_{-\infty}^{\infty} e^{-j2\pi f\tau}\left[\int_{-\infty}^{\infty} x_T(u)x_T(u-\tau)du\right]d\tau$$

除以$2T$得

$$\frac{|X_T(f)|^2}{2T}=\int_{-\infty}^{\infty} e^{-j2\pi f\tau}\left[\frac{1}{2T}\int_{-\infty}^{\infty} x_T(u)x_T(u-\tau)du\right]d\tau$$

取期望和极限

$$G(f)=\lim_{T\to\infty}\left\{E\left[\frac{|X_T(f)|^2}{2T}\right]\right\}$$

$$=\lim_{T\to\infty}\left\{\int_{-\infty}^{\infty} e^{-j2\pi f\tau} E[\hat{R}_T(\tau)]d\tau\right\}$$

其中

$$\hat{R}_T(\tau)=\begin{cases}\dfrac{1}{2T}\displaystyle\int_{-(T-|\tau|)}^{T} x(u)x(u-|\tau|)du & |\tau|\leqslant 2T\\ 0 & |\tau|>2T\end{cases} \tag{2-4-20}$$

故对于$|\tau|\leqslant 2T$有

$$E[\hat{R}_T(\tau)]=\frac{1}{2T}\int_{-(T-|\tau|)}^{T} E[x(u)x(u-|\tau|)]du$$

$$=\frac{1}{2T}\int_{-(T-|\tau|)}^{T} R(\tau)du$$

$$=R(\tau)\left[1-\frac{|\tau|}{2T}\right] \tag{2-4-21}$$

于是

$$G(f)=\lim_{T\to\infty}\left[\int_{-2T}^{2T}\left(1-\frac{|\tau|}{2T}\right)R(\tau)e^{-j2\pi f\tau}d\tau\right] \tag{2-4-22}$$

由于当$T\to\infty$时$\left(1-\frac{|\tau|}{2T}\right)\to 1$，如果当$\tau\to\pm\infty$时$R(\tau)$衰减得足够快从而是绝对可积的，则富氏积分存在，从而证明了下述富氏变换关系：

$$G(f)=\int_{-\infty}^{\infty}R(\tau)e^{-j2\pi f\tau}d\tau=\mathscr{F}[R(\tau)] \tag{2-4-23}$$

$$R(\tau)=\int_{-\infty}^{\infty}G(f)e^{j2\pi f\tau}df=\mathscr{F}^{-1}[G(f)] \tag{2-4-24}$$

令$\tau=0$可得

$$\sigma_x^2=R(0)=\int_{-\infty}^{\infty}G(f)df \tag{2-4-25}$$

σ_x^2是$x(t)$的方差，也代表过程的"总功率"。

归一化的功率谱密度定义为

$$S(f)=G(f)/\sigma_x^2 \tag{2-4-26}$$

故有

$$\rho(\tau)=R(\tau)/R(0)=\int_{-\infty}^{\infty}S(f)e^{j2\pi f\tau}df \tag{2-4-27}$$

$$S(f)=\int_{-\infty}^{\infty}\rho(\tau)e^{-j2\pi f\tau}d\tau \tag{2-4-28}$$

$S(f)$的主要性质如下

(1) $$\int_{-\infty}^{\infty}S(f)df=1 \tag{2-4-29}$$

这由式(2-4-27)中$\tau=0$可得，

(2) $S(f)\geqslant 0$ 　　对任何f 　　(2-4-30)

由于$|X_T(f)|^2\geqslant 0$，故$G(f)\geqslant 0$，因此有这一性质。

(3) 对于实值过程，由$\rho(\tau)=\rho(-\tau)$可得

$S(f)=S(-f)$ 　　任何f 　　(2-4-31)

可能某些平稳过程$G(f)$或$S(f)$不存在(并非对所有f均有限)，但却存在下述积分

$$H(f)=\int_{-\infty}^{f} G(\theta)d(\theta) \tag{2-4-32}$$

或

$$F(f)=\int_{-\infty}^{f} S(\theta)d(\theta) \tag{2-4-33}$$

$H(f)$称为积分功率谱，$F(f)$称归一化积分功率谱。它们和$R(\tau)$或$\rho(\tau)$之间也有一定关系，这种关系首先由维纳，随后由辛钦(Хинчин)所严格证明，现引述如下

维纳-辛钦定理——$\rho(\tau)$作为连续平稳随机过程$\{x(t)\}$的自相关函数的必要和充分条件是：存在一个函数$F(f)$，它在$(-\infty,\ \infty)$范围内具有分布函数的特征(即$F(-\infty)=0, F(+\infty)=1$，且$F(f)$非降）使得在所有$\tau$处$\rho(\tau)$可以表示成

$$\rho(\tau)=\int_{-\infty}^{\infty} e^{j2\pi f\tau}dF(f) \tag{2-4-34}$$

该积分称为富里叶-司蒂吉斯(*Fourier-Stieltjes*)变换，$F(f)$就是归一化积分功率谱，即便F并非处处可微（例如在一些f处有阶跃），式(2-4-34)积分依然存在。如果在所有f处存在$dF(f)/df=S(f)$，从而$dF(f)=S(f)df$，则上式也就是式(2-4-27)。

3. 离散过程的功率谱和自协方差函数的关系 现在转为讨论离散平稳随机过程$\{x_t; t=0, \pm1, \pm2, \cdots\}$的功率谱，它有着和连续过程完全并行的一套关系和类似的物理解释。不同之处有二，(1)因t限为整数，故$R(r)$和$\rho(r)$也只在整数r处有定义，式(2-4-23)和(2-4-28)积分要换成离散和；(2)当t取为相隔为1的整数时，所有谱函数只在$-1/2\leqslant f\leqslant 1/2$范围内有定义。

下面给出离散过程功率谱和自相关函数关系的定理。

沃尔德(*Wold*)定理——序列$\{\rho(r);\ r=0,\ \pm1,\ \pm2, \cdots\}$作为离散平稳随机过程$\{x_t;\ t=0,\ \pm1,\ \pm2,\ \cdots\}$的自相关函

数的必要和充分条件是：存在一个函数$F(f)$，它在$(-\frac{1}{2},\ \frac{1}{2})$区间内具有分布函数的特征（即$F(-\frac{1}{2})=0$，$F(\frac{1}{2})=1$，且$F(f)$非降），使得

$$\rho(r)=\int_{-1/2}^{1/2} e^{j2\pi fr} dF(f) \quad r=0,\pm 1,\pm 2,\cdots \tag{2-4-35}$$

沃尔德定理可以从维纳—辛钦定理延伸而来，为此可以先将$\rho(r)$的离散点用直线逐段连成连续曲线$\rho^c(\tau)$（见图2-13）

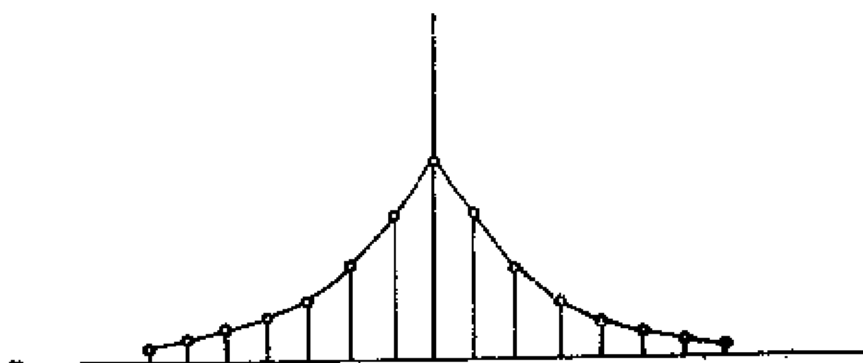


图2-13 $\rho(r)$和$\rho^c(\tau)$图形

$$\rho^c(\tau)=(s+1-\tau)\rho(s)+(\tau-s)\rho(s+1)$$

式中s按离散取值，而$s\leqslant\tau<s+1$。显然$\rho^c(\tau)$处处连续且$\rho^c(-\tau)=\rho^c(\tau)$，同时可以证明它满足作为某一连续平稳过程自相关函数的条件[3]，因此由维纳—辛钦定理，必有一非降函数$Q(f)$使得

$$\rho^c(\tau)=\int_{-\infty}^{\infty} e^{j2\pi f\tau} dQ(f) \tag{2-4-36}$$

由于这一关系对所有τ成立，所以当τ在某些整数r处亦成立，这时$\rho^c(r)$和$\rho(\tau)$等同。

$$\rho(r)=\int_{-\infty}^{\infty} e^{j2\pi fr} dQ(f) \qquad r=0,\pm 1,\pm 2,\cdots \tag{2-4-37}$$

将积分区间分成不重叠的长度为1的小区间，则

$$\rho(r)=\sum_{s=-\infty}^{\infty}\int_{s-1/2}^{s+1/2}e^{j2\pi fr}\ dQ(f)$$

$$=\sum_{s=-\infty}^{\infty}\int_{-1/2}^{1/2}e^{j2\pi(f+s)r}dQ(f+s)$$

由于s,r为整数，$e^{j2\pi sr}=1$，故有

$$\rho(r)=\sum_{s=-\infty}^{\infty}\int_{-1/2}^{1/2}e^{j2\pi fr}\ dQ(f+s)$$

$$=\int_{-1/2}^{1/2}e^{j2\pi fr}\left[\sum_{s=-\infty}^{\infty}dQ(f+s)\right]$$

令

$$dF(f)=\sum_{s=-\infty}^{\infty}dQ(f+s),\qquad 且F(-\tfrac{1}{2})=0$$

则由$Q(f)$的性质可证明

$$F(\tfrac{1}{2})=\int_{-1/2}^{1/2}\left[\sum_{s=-\infty}^{\infty}dQ(f+s)\right]=\int_{-\infty}^{\infty}dQ(f)=Q(\infty)=1$$

这就得出沃尔德定理所给出的式(2-4-35)

从上面讨论可以看出，当t限为整数值(r)时，我们无法区分$e^{j2\pi fr}$和$e^{j2\pi(f+s)r}$(s为任意整数)。因此，如果离散过程$\{x_t\}$是连续过程$\{x(t)\}$在$t=0,\pm1,\pm2,\cdots$处的采样，则$x(t)$中所有频率为$f\pm1$，$f\pm2$，…的分量都以f的频率包含在$F(f)$中，这就是离散采样导致的“频率混叠”。由于$F(f)$的频率范围在$(-\frac{1}{2},\frac{1}{2})$内，故在这范围以外的频率成份都被纳入其中。

如果所论$\{x_t\}$的$F(f)$是处处可微的，则归一化功率谱$S(f)=dF(f)/df$存在，此时式(2-4-35)可写成

$$\rho(r)=\int_{-1/2}^{1/2}S(f)e^{j2\pi fr}\ df\quad r=0,\pm1,\pm2\cdots\qquad(2\text{-}4\text{-}38)$$

将该式和式(2-4-3)(2-4-4)比较，并考虑到$S(f)$的周期为1，$\rho(r)$是$S(f)$复富氏级数展开式中$e^{-j2\pi fr}$的系数，因此式(2-4-38)的反演为

$$S(f)=\sum_{r=-\infty}^{\infty}\rho(r)e^{-j2\pi fr}\qquad -\frac{1}{2}\leqslant f\leqslant\frac{1}{2}\qquad(2\text{-}4\text{-}39)$$

若$\{x_t\}$为实序列，则$\{\rho(r)\}$为偶序列，式(2-4-39)简化为余弦变换

$$S(f)=\sum_{r=-\infty}^{\infty}\rho(r)\cos 2\pi fr$$

$$=1+2\sum_{r=1}^{\infty}\rho(r)\cos 2\pi fr\qquad -\frac{1}{2}\leqslant f\leqslant\frac{1}{2}\qquad(2\text{-}4\text{-}40)$$

和式(2-4-38)～(2-4-40)相应，非归一化的情况下有

$$R(r)=\int_{-1/2}^{1/2}e^{j2\pi fr}\,G(f)df\qquad r=0,\pm1,\pm2,\cdots\qquad(2\text{-}4\text{-}41)$$

$$G(f)=\sum_{r=-\infty}^{\infty}R(r)e^{-j2\pi fr}\qquad -\frac{1}{2}\leqslant f\leqslant\frac{1}{2}\qquad(2\text{-}4\text{-}42)$$

过程为实值时

$$G(f)=\sum_{r=-\infty}^{\infty}R(r)\cos 2\pi fr=\sigma_x^2+2\sum_{r=1}^{\infty}R(r)\cos 2\pi fr$$

$$-\frac{1}{2}\leqslant f\leqslant\frac{1}{2}\qquad(2\text{-}4\text{-}43)$$

不难看出，方差为σ_x^2的白噪声过程的功率密度谱为

$$G(f)=\sigma_x^2\qquad -\frac{1}{2}\leqslant f\leqslant\frac{1}{2}\qquad(2\text{-}4\text{-}44)$$

$$S(f)=1\qquad -\frac{1}{2}\leqslant f\leqslant\frac{1}{2}\qquad(2\text{-}4\text{-}45)$$

4. **基本线性模型的功率谱** 根据自相关函数和功率谱密度的关系，可以导出线性模型的功率谱密度函数。

对式(2-3-73)式两边取富氏变换得

$$G_x(f)=\sum_{\tau=-\infty}^{\infty}\sum_{i=0}^{\infty}G_i\sum_{l=0}^{\infty}G_l R_\varepsilon(\tau+i-l)e^{-j2\pi f\tau}$$

$$=\sum_{i=0}^{\infty}G_i e^{j2\pi fi}\sum_{l=0}^{\infty}G_l e^{-j2\pi fl}\sum_{\tau=-\infty}^{\infty}R_\varepsilon(\tau+i-l)$$

$$e^{-j2\pi f(\tau+i-l)}$$

该式第二行的头两个和式可看作（见式(2-3-35)）算子式中以$e^{-j2\pi f}$代替B的结果，即

$$G(B)|_{B=\exp(-j2\pi f)}=[\Phi^{-1}(B)\Theta(B)]_{B=\exp(-j2\pi f)}=G(f)$$

$$G(B^{-1})|_{B=\exp(-j2\pi f)}=[\Phi^{-1}(B^{-1})\Theta(B^{-1})]_{B=\exp(-j2\pi f)}$$

$$=G^*(f)$$

因此

$$G_x(f)=G^*(f)G(f)G_\varepsilon(f)=|G(f)|^2G_\varepsilon(f)$$

$$=|G(B)|^2_{B=\exp(-j2\pi f)}G_\varepsilon(f)$$

$$=\left|\frac{\Theta(B)}{\Phi(B)}\right|^2_{B=\exp(-j2\pi f)}G_\varepsilon(f)\quad -\tfrac{1}{2}\leqslant f\leqslant\tfrac{1}{2}\quad(2\text{-}4\text{-}46)$$

在ε_t是方差为σ_ε^2的白噪声的情况下

$$G_x(f)=\sigma_\varepsilon^2\left|\frac{\Theta(B)}{\Phi(B)}\right|^2_{B=\exp(-j2\pi f)}\quad -\tfrac{1}{2}\leqslant f\leqslant\tfrac{1}{2}\quad(2\text{-}4\text{-}47)$$

由于Θ和Φ是两个有限阶的多项式，因此$\left|\frac{\Theta(B)}{\Phi(B)}\right|_{B=\exp(-j2\pi f)}$是两个$\exp(-j2\pi f)$的多项式之比，它是$\exp(-j2\pi f)$的有理函数，这就是ARMA过程谱密度的形式。

在工程上常把上述结论表达成：ARMA过程是由白噪声过程通过一个有理传递函数$\Theta(B)/\Phi(B)$的滤波器所产生的输出(见图2-14)

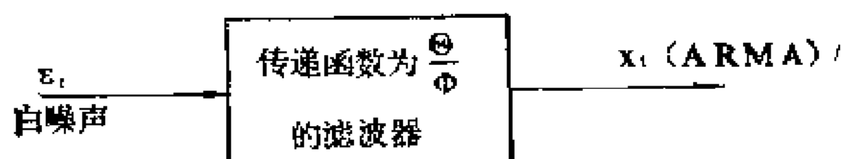


图2-14　过程的传递函数模型图

式(2-4-46)在研究线性变换中是很基本的，它表明输出的非归一化谱密度函数等于输入的非归一化谱密度乘以传递函数模的平方。

值得指出的是：输出功率谱在频率f处的值只依赖于同一频率f处的输入功率谱和$G(f)$的值，而不受输入功率谱在其他频率处的值的影响。这一点和时间域中输入和输出的关系是截然不同的。

利用式(2-4-47)可以很容易地根据过程的差分方程写出相应的功率谱表达式。

AR(1)过程$x_t+ax_{t-1}=\varepsilon_t$的功率谱为

$$G_x(f)=\sigma_\varepsilon^2\frac{1}{|1+aB|^2}\bigg|_{B=\exp(-j2\pi f)}=\sigma_\varepsilon^2\frac{1}{|1+a\exp(-j2\pi f)|^2}$$

$$=\sigma_\varepsilon^2/(1+a^2+2a\cos 2\pi f)\qquad -\tfrac{1}{2}\leqslant f\leqslant\tfrac{1}{2}\qquad(2\text{-}4\text{-}48)$$

图2-15为$\sigma_\varepsilon^2=1$，和不同a值时的$G_x(f)$。a为负时主要功率集中在低频部分，而a为正时则集中在高频部分。

如果考虑到$\sigma_x^2=\sigma_\varepsilon^2/(1-a^2)$，则$AR(1)$的归一化功率谱

$$S_x(f)=\frac{1-a^2}{1+a^2+2a\cos 2\pi f}\qquad -\tfrac{1}{2}\leqslant f\leqslant\tfrac{1}{2}\qquad(2\text{-}4\text{-}49)$$

$AR(2)$过程$x_t+a_1x_{t-1}+a_2x_{t-2}=\varepsilon_t$的功率谱为

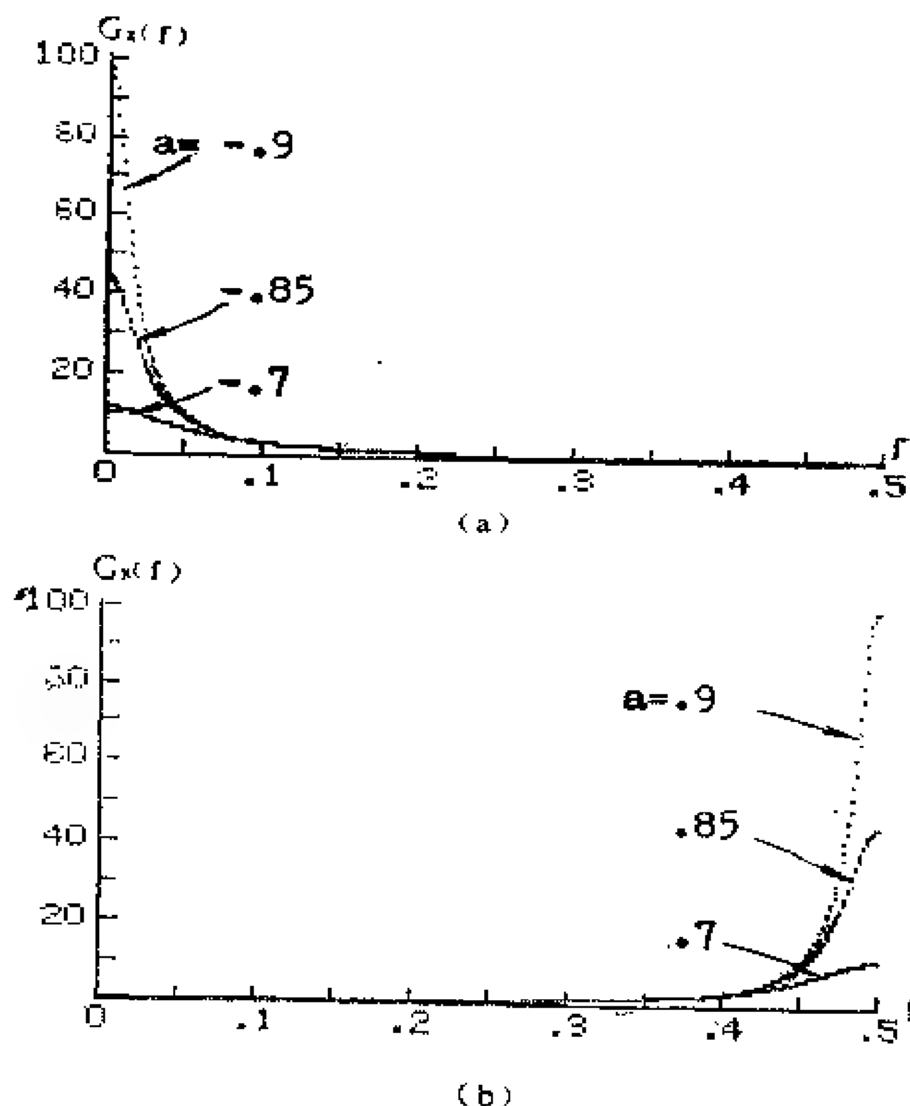


图2-15　AR(1)过程的功率谱

$$G_x(f)=\sigma_x^2 S_x(f)=\frac{\sigma_e^2}{|1+a_1\exp(-j2\pi f)+a_2\exp(-j4\pi f)|^2}$$

$$=\frac{\sigma_e^2}{1+a_1^2+a_2^2+2a_1(1+a_2)\cos 2\pi f+2a_2\cos 4\pi f}$$

$$-\frac{1}{2}\leqslant f\leqslant\frac{1}{2} \qquad (2\text{-}4\text{-}50)$$

当$a_2>0$，且$|a_1(1+a_2)|<|4a_2|$时，$G_x(f)$会出现峰值，与此峰值对应的频率是f_0，而

$$\cos 2\pi f_0 = -a_1(1+a_2)/(4a_2)$$

图2-16是$a_1=-1$，$a_2=0.5$时的$G_x(f)$曲线.

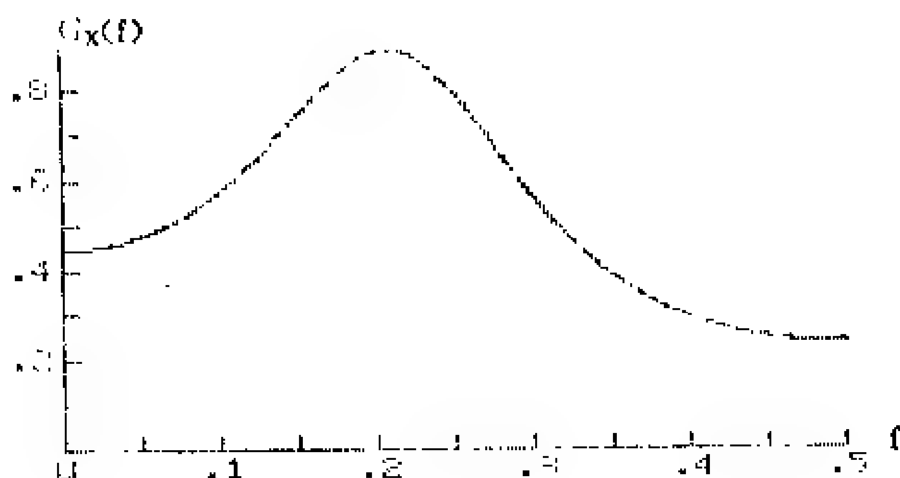


图2-16　AR(2)过程的功率谱

若将式(2-3-38)代入式(2-4-50)则有

$$S_x(f)=\frac{(1-a_2)[(1+a_2)^2-a_1^2]}{(1+a_2)[1+a_1^2+a_2^2+2a_1(1+a_2)\cos 2\pi f+2a_2\cos 4\pi f]}$$

$$-\tfrac{1}{2}\leqslant f\leqslant \tfrac{1}{2} \qquad (2\text{-}4\text{-}51)$$

AR(n)过程$x_t+a_1x_{t-1}+\cdots+a_nx_{t-n}=\varepsilon_t$的功率谱

$$G_x(f)=\sigma_x^2S_x(f)=\sigma_\varepsilon^2/|1+a_1\exp(-j2\pi f)+\cdots+a_n\exp(-j2\pi nf)|^2$$

$$-\tfrac{1}{2}\leqslant f\leqslant \tfrac{1}{2} \qquad (2\text{-}4\text{-}52)$$

MA(m)过程$x_t=b_0\varepsilon_t+b_1\varepsilon_{t-1}+\cdots+b_m\varepsilon_{t-m}$的功率谱

$$G_x(f)=\sigma_x^2S_x(f)=\sigma_\varepsilon^2|b_0+b_1\exp(-j2\pi f)+\cdots+b_m\exp(-j2\pi mf)|^2$$

$$-\tfrac{1}{2}\leqslant f\leqslant \tfrac{1}{2} \qquad (2\text{-}4\text{-}53)$$

在$b_0=b_1=\cdots=b_m=1/(m+1)$的特殊情况下

$$G_x(f)=\sigma_\varepsilon^2[\sin^2(m+1)\pi f]/[(m+1)^2\sin^2\pi f]$$

$$-\tfrac{1}{2}\leqslant f\leqslant \tfrac{1}{2} \qquad (2\text{-}4\text{-}54)$$

ARMA(n,m)过程$x_t+a_1x_{t-1}+\cdots+a_nx_{t-n}=b_0\varepsilon_t+b_1\varepsilon_{t-1}$

$+\cdots+b_m\varepsilon_{t-m}$的功率谱为

$$G_x(f)=\sigma_x^2 S_x(f)=\sigma_\varepsilon^2\left|\frac{b_0+b_1\exp(-j2\pi f)+\cdots+b_m\exp(-j2\pi mf)}{1+a_1\exp(-j2\pi f)+\cdots+a_n\exp(-j2\pi nf)}\right|^2$$

$$-\frac{1}{2}\leqslant f\leqslant\frac{1}{2} \qquad (2\text{-}4\text{-}55)$$

应当注意，如果根据给定的$G_x(f)$确定ARMA模型参数，则应考虑到构成$|\Phi(B)|^2$的$2n$阶多项式中B的根在B平面单位圆内外各占半数且互为倒数，为保证平稳，应取在单位圆外的根构成$\Phi(B)$。$|\Theta(B)|^2$中根的分布也类似，在全部根中可有2^m种不同的组合构成$\Theta(B)$而都具有相同的$|\Theta(B)|^2_{B=\exp(-j2\pi f)}$，但其中只有一种能使$\Theta(B)$的根符合可逆性条件。因此在给定$G_x$（或$R_x$）时，平稳和可逆条件保证了模型参数的唯一性。

5. 平稳过程的谱展式 在2.4.1节中曾指出，不论富氏级数还是富氏积分都不能将一个平稳随机过程直接分解，当时并没有讨论是否有别的富氏型式的展开可以做到这一点，而是绕过这一困难，先用有限区间内一个实现的富氏积分来表示，然后考虑区间增大的极限情况来定义过程的功率谱密度。究竟能否把一个完全的实现$(-\infty<t<\infty)$表示成sin和cos项之和？做到这一点，才能断言是"$x(t)$中某频率f的分量"对总功率的贡献，因为所谓"频率"的含意实质上暗示着过程中存在有sin和cos信号。维纳提出的"广义谐波分析"严格证明了用富里叶—司蒂吉斯积分可以统一地代表富氏级数和通常的富氏积分，而且还可表示那些不具备富氏级数或富氏积分条件的函数。

下面先给出谱展式的定理，然后加以启发式的说明。

连续平稳过程的谱展式定理——设$\{x(t)\}$，$-\infty<t<\infty$为零均值连续平稳随机过程，则有一正交过程$\{Z(f)\}$可将所有t下的$x(t)$表示成下列均方意义下的积分，即

$$x(t)=\int_{-\infty}^{\infty} e^{j2\pi ft}dZ(f) \tag{2-4-56}$$

其中过程 $\{Z(f)\}$ 具有如下性质：

(1) $E[dZ(f)]=0$ 所有f

(2) $E[|dZ(f)|^2]=dH(f)$ 所有f (2-4-57)

这里$H(f)$是$x(t)$的非归一化积分谱

(3)在不同频率f和$f'(f\neq f')$处

$$\mathrm{Cov}[dZ(f),\ dZ(f')]=E[dZ^*(f)dZ(f')]=0 \tag{2-4-58}$$

式(2-4-56)所示富里叶—司蒂吉斯积分称为$x(t)$过程的谱展式，下面进行推导并证明其性质。

先取有限区间$(-T,T)$内的一个实现$x(t)$，并使其在区间外为周期重复，即定义

$$\begin{aligned} x_{T\bar{p}}(t)&=x(t) & -T\leqslant t\leqslant T \\ x_{T\bar{p}}(t+2nT)&=x_{T\bar{p}}(t) & n=\pm1,\pm2,\cdots \end{aligned} \tag{2-4-59}$$

用富氏级数表达成

$$x_{T\bar{p}}(t)=\sum_{k=-\infty}^{\infty} A_k e^{j2\pi f_k t} \tag{2-4-60}$$

其中$f_k=k/(2T)$，而

$$A_k=\frac{1}{2T}\int_{-T}^{T} x_{T\bar{p}}(t)e^{-j2\pi f_k t}\,dt \tag{2-4-61}$$

定义

$$X_T(f)=\int_{-T}^{T} x_{T\bar{p}}(t)e^{-j2\pi ft}dt=\int_{-T}^{T} x(t)e^{-j2\pi ft}dt \tag{2-4-62}$$

注意到$1/(2T)=f_{k+1}-f_k=\delta f_k$，式(2-4-60)可写成

$$x_{T\bar{p}}(t)=\sum_{k=-\infty}^{\infty} X_T(f_k)e^{j2\pi f_k t}\delta f_k \tag{2-4-63}$$

尽管$x_{T\bar{p}}(t)$和式(2-4-15)定义的$x_T(t)$不同，但式（2-4-62）的$X_T(f)$和式(2-4-17)定义的$X_T(f)$是相同的。已曾指出，当谱密度$G(f)$存在时有极限

$$\lim_{T\to\infty}\left\{E\left[\frac{|X_T(f)|^2}{2T}\right]\right\}=G(f)$$

因此$T\to\infty$时$|X_T(f_k)|=O(\sqrt{T})=O(1/\sqrt{\delta f})$，可见$T\to\infty$时$|X_T(f_k)|\to\infty$，所以不能在式(2-4-63)中通过$T\to\infty$来得到$x(t)$的展开式，因为$X_T(f_k)$不能收敛到有限值。若考虑$X_T(f_k)\delta f_k$则不同，因为$|X_T(f_k)|\delta f_k$-$O(\sqrt{\delta f})$当$T\to\infty$时趋于零。定义函数

$$Z_T(f)=\int_{-\infty}^{f} X_T(\theta)d\theta \tag{2-4-64}$$

则T很大或δf_k很小时

$$\Delta Z_T(f_k)=Z_T(f_{k+1})-Z_T(f_k)\simeq X_T(f_k)\delta f_k \tag{2-4-65}$$

于是

$$x_{T\bar{p}}(t)=\sum_{k=-\infty}^{\infty} e^{j2\pi f_k t}\Delta Z_T(f_k) \tag{2-4-66}$$

令$T\to\infty$，从而$\delta f_k\to 0$（即f_k的间隔趋于零），式(2-4-66)右边收敛于适合任何t的积分，即

$$x(t)=\int_{-\infty}^{\infty} e^{j2\pi ft}\ dZ(f)$$

上面只提到一个实现，如果要用上式表示各个实现，则$Z(f)$对所有的实现而言将是不同的，或者说$Z(f)$本身是一个以f为变量的复随机过程（对每个f，$Z(f)$是复随机变量），这时上式右边是均方意义下的随机积分。这个积分形式上和式(2-4-34)相同，但$Z(f)$和积分功率谱$F(f)$性质不同，$F(f)$是可微的或阶跃的，当它可微时式(2-4-34)可化为通常的富氏积分，但$F(f)$可微时$Z(f)$

并不可微，因为$dZ(f)=O(\sqrt{df})$，所以"导数"$dZ(f)/df$不存在。

$x(t)$的富里叶-司蒂吉斯积分表明了可以把$x(t)$表示成具有随机系数$dZ(f)$的正弦和余弦函数之和，或者说可分解成具有随机振幅$|dZ(f)|$和随机相位$\arg[dZ(f)]$的周期函数之和。

由于很难用具体形式表达不同频率下的$dZ(f)$，并考虑到主要关心的是$x(t)$的能量分布，故考察$|dZ(f)|^2$，可证明，它和df是同阶的。事实上，由式(2-4-65)可得

$$E[|\Delta Z_T(f_k)|^2]\simeq E[|X_T(f_k)|^2\delta f_k]\delta f_k=E\left[\frac{|X_T(f_k)|^2}{2T}\right]\delta f_k$$

当$T\to\infty$（即$\delta f_k\to 0$）时

$$E[|\Delta Z(f_k)|^2]\simeq G(f_k)\delta f_k=\Delta H(f_k) \tag{2-4-67}$$

其中　$H(f)=\int_{-\infty}^{f} G(\theta)d\theta$是$x(t)$的非归一化积分功率谱。式(2-4-67)的极限形式为

$$E[|dZ(f)|^2]=G(f)df=dH(f)=\sigma_x^2 dF(f) \tag{2-4-68}$$

其中$F(f)$是$x(t)$的归一化积分谱。如果对任何f均存在谱密度$G(f)$和$S(f)$，则

$$E[|dZ(f)|^2]/df=G(f)=\sigma_x^2 S(f) \tag{2-4-69}$$

现在来证明$\{Z(f)\}$为正交过程，这是指在不同f处Z的增量是无关的，即若$f\neq f'$，则随机变量$dZ(f)=[Z(f+df)-Z(f)]$和$dZ(f')=[Z(f'+df')-Z(f')]$的协方差为零。

首先，因$x(t)$均值为零，由式(2-4-62)得$E[X_T(f)]=0$（所有f），从而$E[dZ(f)]=0$，在式(2-4-56)两边取共轭并乘以$x(t+\tau)$得

$$x^*(t)x(t+\tau)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-j2\pi ft}e^{j2\pi f'(t+\tau)}dZ^*(f)dZ(f')$$

再取期望

$$E[x^*(t)x(t+\tau)]$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{j2\pi(f'-f)t}e^{j2\pi f'\tau}E[dZ^*(f)dZ(f')]$$

$$(2\text{-}4\text{-}70)$$

这是平稳过程的自协方差函数，应与t无关，因此$f\neq f'$时上式右边二重积分必为零，从而有

$$E[dZ^*(f)dZ(f')]=0 \qquad f\neq f'$$

$Z(f)$的正交性质得证，根据这一性质，式(2-4-70)可化为单积分，即取$f=f'$有

$$R(\tau)=E[x^*(t)x(t+\tau)]=\int_{-\infty}^{\infty} e^{j2\pi f\tau}\ E[|dZ(f)|^2]$$

$$=\int_{-\infty}^{\infty} e^{j2\pi f\tau}\ dH(f) \qquad (2\text{-}4\text{-}71)$$

这也正是维纳-辛钦定理的非归一化形式。

对于离散平稳序列的谱展式，只须对连续情况的结论作些修改即可，这就是t取整数值并将式(2-4-56)的积分限改为$-\frac{1}{2}$到$\frac{1}{2}$。定理如下：

若$\{x_t\}$ $t=0,\ \pm1,\ \pm2,\ \cdots$为零均值平稳过程，则在区间$(-\frac{1}{2},\frac{1}{2})$内有一正交过程$\{Z(f)\}$使得所有整数$t$处

$$x_t=\int_{-\frac{1}{2}}^{\frac{1}{2}} e^{j2\pi ft}\ dZ(f) \qquad (2\text{-}4\text{-}72)$$

其中$Z(f)$的性质和连续过程相同，但

$$E[|dZ(f)|^2]=dH(f) \qquad -\tfrac{1}{2}\leqslant f\leqslant \tfrac{1}{2} \qquad (2\text{-}4\text{-}73)$$

$x(t)$或x_t的谱展式定理在整个平稳过程理论中是非常重要的，有关结论在多变量过程（第七章）中还要多次用到。

第三章　时间域模型的估计

上一章我们讨论了线性平稳时间序列的性质，建立了表征过程性质的自协方差（自相关）和功率密度谱等非参数模型，也讲述了可以用有限个参数表示的ARMA模型（称参数模型）。由于我们实际掌握的是时间序列的一些原始观察数据，往往不能事先确认它们完全符合线性模型，同时也不能单纯按机理直接推导的办法建立它们的模型．因而需要一种实际可行的估计模型的方法，这就是进行相关函数估计（属于非参数估计），及采用时间域的拟合方法来获得尽可能与实际过程等价的ARMA模型。

根据已掌握的一组样本数据序列x_1，x_2，…，x_N建立ARMA模型，其含义就是对模型的阶数（n，m）和参数（包括μ，a_i，b_i和σ_a^2）作出判断或估计。一般称n，m的判断为定阶，而估计相应的μ、a_i、b_i、σ_a^2称为参数估计。

3.1　自协方差和自相关函数的估计

1．两种估计方法　设$\{x_t\}$为一离散平稳过程，它具有

均值$E[x_t]=\mu$　　　　（对任何t）

方差$\mathrm{Var}[x_t]=\sigma^2$　　　　（对任何t）

自协方差$\mathrm{Cov}[x_t,\ x_{t+r}]=R(r)=\sigma^2\rho(r)$

这些μ，σ^2，$R(r)$或$\rho(r)$均属未知，已知的只是N个观察值x_1，x_2，…，x_N，主要的任务是估计$R(r)$或$\rho(r)$，但是首先要估计μ。

对于平稳过程而言，x_1，x_2，…，x_N都具有相同的均值μ，因此有理由用样本的均值作为μ的估值，即

$$\bar{x}=\frac{1}{N}\sum_{t=1}^{N}x_t \tag{3-1-1}$$

由于$E(\bar{x})=\frac{1}{N}\sum_{t=1}^{N}E(x_t)=\frac{1}{N}(N\mu)=\mu$，可见$\bar{x}$是$\mu$的无偏估计。

均值估计的方差为

$$\begin{aligned}\mathrm{Var}(\bar{x})&=E\left(\frac{1}{N}\sum_{t=1}^{N}x_t-\mu\right)^2=E\left(\frac{1}{N}\sum_{t=1}^{N}(x_t-\mu)\right)^2\\&=\frac{1}{N^2}\sum_{s=1}^{N}\sum_{t=1}^{N}E[(x_s-\mu)(x_t-\mu)]=\frac{\sigma^2}{N^2}\sum_{s=1}^{N}\sum_{t=1}^{N}\rho(t-s)\end{aligned}$$

令$t-s=r$，将上式二重求和化作对r求和，并考虑到$\rho(r)$为偶函数，可得

$$\mathrm{Var}(\bar{x})=\frac{\sigma^2}{N}\sum_{r=-(N-1)}^{N-1}\left(1-\frac{|r|}{N}\right)\rho(r)$$

当N很大时，和式近似为 $\sum_{r=-\infty}^{\infty}\rho(r)=S(f)\Big|_{f=0}=S(0)$

故

$$\mathrm{Var}(\bar{x})\cong\frac{\sigma^2}{N}S(0)$$

由于估计值$\bar{x}$的方差是$O(1/N)$，因而$\bar{x}$是μ的一致估计。

自协方差的估计则是从一组观察值x_1，…，x_N中取相隔为r的各对观察值来求平均，一共有$N-r$对：

$(x_1,x_{r+1}),(x_2,x_{r+2})$，…，$(x_{N-r},\ x_N)$

显然，$\{x_1,\ \cdots,\ x_{N-r}\}$和$\{x_{r+1},\ \cdots x_N\}$都具有相同的均值$\mu$，因此对每个$r$有估计值

$$\hat{R}^{(*)}(r)=\frac{1}{N-|r|}\sum_{t=1}^{N-|r|}(x_t-\bar{x})(x_{t+|r|}-\bar{x}) \tag{3-1-2}$$

$$r=0,\ \pm 1,\ \cdots,\ \pm(N-1)$$

（这里附加(＊)号是为了和下面的另一种$R(r)$的估计相区别）

注意：由于只有N个采样点，即不存在间隔大于$(N-1)$的观察值，所以不可能估计$|r|\geqslant N$的$R(r)$值。此外，由于$R(r)$是偶函数，对所有的r存在$R(-r)=R(r)$，因而$\hat{R}^{(*)}(r)$也应当是偶函数，为此在式中取$|r|$。

下面我们来考察$\hat{R}^{(*)}(r)$的无偏性。

如果式(3.1.2)中的$\bar{x}$是真正的均值μ，则

$$E[\hat{R}^{(*)}(r)]=\frac{1}{N-|r|}\sum_{t=1}^{N-|r|}E[(x_t-\mu)(x_{t+r}-\mu)]$$

$$=\frac{1}{N-|r|}\sum_{t=1}^{N-|r|}R(r)=R(r) \qquad (3\text{-}1\text{-}3)$$

然而实际上

$$\sum_{t=1}^{N-|r|}[(x_t-\mu)(x_{t+|r|}-\mu)]=\sum_{t=1}^{N-|r|}\{(x_t-\bar{x})+(\bar{x}-\mu)\}$$

$$\{(x_{t+|r|}-\bar{x})+(\bar{x}-\mu)\}\cong\sum_{t=1}^{N-|r|}[(x_t-\bar{x})(x_{t+|r|}-\bar{x})]$$

$$+(N-|r|)\ (\bar{x}-\mu)^2$$

两边取数学期望得

$$R(r)=E[\hat{R}^{(*)}(r)]=E(\bar{x}-\mu)^2]$$

因此

$$E[\hat{R}^{(*)}(r)]=R(r)-E[\bar{X}-\mu)^2]=R(r)-Var(\bar{x})$$

$$=R(r)-\frac{\sigma^2}{N}S(0) \qquad (3\text{-}1\text{-}4)$$

换言之，$\hat{R}^{(*)}$是有偏估计，其偏差是$O(\frac{1}{N})$，当然也是渐近无偏的

$R(r)$的另一种估计式是以N作为除数（而不是$N-|r|$）：

$$\hat{R}(r)=\frac{1}{N}\sum_{t=1}^{N-|r|}(x_t-\bar{x})(x_{t+|r|}-\bar{x}) \tag{3-1-5}$$

可以证明，当忽略均值偏差的情况下（即认为$\bar{x}=\mu$）有

$$E[\hat{R}(r)]=\frac{1}{N}\sum_{t=1}^{N-|r|}R(r)=(1-\frac{|r|}{N})R(r) \tag{3-1-6}$$

这是一个有偏估计。可见，当r为某定值时该估计的偏差是$O(\frac{1}{N})$，但偏差随着r增大而增大。如果再考虑均值估计的偏差。则有

$$E[\hat{R}(r)]=R(r)-\frac{|r|}{N}R(r)-\frac{N-|r|}{N^2}\sigma^2S(0) \tag{3-1-7}$$

在不考虑均值估计偏差的情况下，$\hat{R}(r)$的数学期望为

$$E[\hat{R}(r)]=(1-\frac{|r|}{N})R(r) \tag{3-1-8}$$

它取决于$|r|$较N的相对大小以及$R(r)$值的大小。当$|r|$比N相对较小时，$E[\hat{R}(r)]$对$R(r)$的偏差小，当$|r|$接近$(N-1)$时偏差实际上就是$R(r)$。不过对不含周期成份的过程而言，当$r\to\infty$时，$R(r)\to0$，因此只要N较大（相对于$R(r)$的衰减快慢而言），则当$|r|$趋近于$(N-1)$时，$R(r)$本身就很小，所以在整个r的范围内，$\hat{R}(r)$对$R(r)$的平均偏差都很小。

通常粗略地称$\hat{R}^{(*)}(r)$为无偏估计，$\hat{R}(r)$为有偏估计。虽然后者是有偏的，但它的均方差却比前者小，根据巴特利公式（见下节）可以导出

$$\mathrm{Var}[\hat{R}(r)]=O(\frac{1}{N}) \quad 对所有r \tag{3-1-9}$$

而

$$\mathrm{Var}(\hat{R}^{(*)}(r)) = O(\frac{1}{N-|r|}) \qquad (3\text{-}1\text{-}10)$$

可见当r接近$(N-1)$时，$\mathrm{Var}(\hat{R}^{(*)}(r))$增大很多，而$\hat{R}(r)$的方差对任何$r$均保持$O(\frac{1}{N})$，所以$\hat{R}(r)$的波动在整个$r$的范围内都保持很小。当然，在$|r|$较小时，两种估计相差不大，而当$r$较大时，$\hat{R}(r)$的方差较小，从而大大弥补了它较大的偏差。

现在大多数时间序列分析所使用的是有偏估计$\hat{R}(r)$，因为主要关心的往往不是某一个r值时$R(r)$的估计值，而是从0到$\pm(N-1)$范围内所有r值的函数$R(r)$的估计。此外，$\hat{R}(r)$还具有自协方差函数的半正定特性，而$\hat{R}^{(*)}(r)$却不一定具备。

2．关于协方差和自相关估计的巴特利公式 首先讨论$\hat{R}(r)$的协方差

设μ已知，故不失一般性令$E[x_t]=0$，则有

$$\hat{R}(r) = \frac{1}{N}\sum_{t=1}^{N-|r|} x_t x_{t+|r|} \qquad (3\text{-}1\text{-}11)$$

$$E(\hat{R}(r)) = (1-|r|/N)R(r) \quad 所有r \quad (3\text{-}1\text{-}12)$$

$$\mathrm{Cov}(\hat{R}(r),\hat{R}(r+v)) = E(\hat{R}(r)(r+v)) - E(\hat{R}(r))E(\hat{R}(r+v))$$

$$= \frac{1}{N^2}E(\sum_{t=1}^{N-r}\sum_{s=1}^{N-r-v} x_t x_{t+r} x_s x_{s+r+v}) - (1-\frac{r}{N})(1-\frac{r+v}{N})$$

$$R(r)R(r+v)$$

如果$\{x_t\}$为高斯过程则

$$E[x_t x_{t+r} x_s x_{s+r+v}]$$

$$= E[x_t x_{t+r}]E[x_s x_{s+r+v}] + E[x_t x_s]E[x_{t+r} x_{s+r+v}]$$

$$+ E[x_t x_{s+r+v}]E[x_{t+r} x_s]$$

$$= R(r)R(r+v) + R(s-t)R(s+v-t) + R(s+r+v-t)$$

$$R(s-t-r) \qquad (3\text{-}1\text{-}14)\mathrm{a}$$

$$\mathrm{Cov}[\hat{R}(r),\hat{R}(r+v)]$$

$$=\frac{1}{N^2}\sum_{t=1}^{N-r}\sum_{s=1}^{N-r-v}\{R(s-t)R(s+v-t)+R(s+r+v-t)R(s-t-r)\}$$

$$=\frac{1}{N}\sum_{m=-(N-r)+1}^{N-r-v-1}\left\{1-\frac{\eta(m)+r+v}{N}\right\}\{R(m)R(m+v)$$

$$+R(m+r+v)R(m-r)\} \qquad (3\text{-}1\text{-}14)\text{b}$$

其中

$$\eta(m)=\begin{cases} m & m>0 \\ o & -v\leqslant m\leqslant 0 \\ -m-v & -(N-r)+1\leqslant m<-v \end{cases}$$

此式由巴特利1946年给出。

令$v=0$得

$$\mathrm{Var}[\hat{R}(r)]=\frac{1}{N}\sum_{m=-(N-r)+1}^{N-r-1}\left\{1-\frac{|m|+r}{\mathrm{N}}\right\}$$

$$\{R^2(m)+R(m+r)R(m-r)\} \qquad (3\text{-}1\text{-}15)$$

当N很大时，以上二式可简化成

$$\mathrm{Cov}[\hat{R}(r),\hat{R}(r+v)]\cong\frac{1}{N}\sum_{m=-\infty}^{\infty}\{R(m)R(m+v)$$

$$+R(m+r+v)R(m-r)\} \qquad (3\text{-}1\text{-}16)$$

$$\mathrm{Var}[\hat{R}(r)]\cong\frac{1}{N}\sum_{m=-\infty}^{\infty}\{R^2(m)+R(m+r)R(m-r)\} \qquad (3\text{-}1\text{-}17)$$

以上各式稍加修改便可用于讨论无偏估计，因

$$\hat{R}^{(*)}(r)=\frac{N}{N-r}\hat{R}(r) \qquad (r\geqslant 0) \qquad (3\text{-}1\text{-}18)$$

故

$$\mathrm{Var}[R^{(*)}(r)]\cong\frac{1}{N-r}\sum_{m=-\infty}^{\infty}\{R^2(m)$$

$$+R(m+r)R(m-r)\} \qquad (3\text{-}1\text{-}19)$$

这就说明了$\mathrm{Var}[\hat{R}(r)]=O(\frac{1}{N})$及$\mathrm{Var}[\hat{R}^{(*)}(r)]=O\left(\frac{1}{N-r}\right)$

现在来看自相关函数$\rho(r)=R(r)/R(0)$，取

$$\hat{\rho}(r)=\hat{R}(r)/\hat{R}(0) \quad r=0,\pm 1,\cdots,\pm(N-1) \qquad (3-1-20)$$

以及对应于无偏估计的

$$\hat{\rho}^{(*)}(r)=\hat{R}^{(*)}(r)/\hat{R}(0) \qquad (3-1-21)$$

$\hat{\rho}^{(*)}(r)$和$\hat{\rho}(r)$的差别同$\hat{R}^{(*)}$和$\hat{R}$的差别一样。下面对$\hat{\rho}(r)$进行讨论。用$\delta\hat{R}$表示$\hat{R}$对于其均值的差，即$\delta[\hat{R}(\cdot)]=\hat{R}(\cdot)-E[\hat{R}(\cdot)]$，并假定$\delta[\hat{R}(\cdot)]$比$E[\hat{R}(\cdot)]$小，从而有近似关系

$$E[\hat{\rho}(r)]\simeq E[\hat{R}(r)]/E[\hat{R}(0)]=(1-\frac{|r|}{N})\rho(r) \qquad (3-1-22)$$

这是$\rho(r)$估计的数学期望。下面求$\rho(r)$估计的自协方差。令

$\delta[\hat{\rho}(r)]$表示$\hat{\rho}(r)-E[\hat{\rho}(r)]$，取其一次近似有

$$\delta[\hat{\rho}(r)]=\delta\left[\frac{\hat{R}(r)}{\hat{R}(o)}\right]\cong\frac{\delta[\hat{R}(r)]}{\hat{R}(o)}-\frac{\hat{R}(r)\delta[\hat{R}(o)]}{\hat{R}^2(o)}$$

$$\cong\frac{\delta[\hat{R}(r)]}{R(o)}-\frac{R(r)\delta[\hat{R}(o)]}{R^2(o)}$$

从而有

$$\mathrm{Cov}[\hat{\rho}(r),\hat{\rho}(r+v)]=E(\delta[\hat{\rho}(r)]\delta(\hat{\rho}(r+u)])$$

$$\cong E\left\{\frac{\delta[\hat{R}(r)]\delta[\hat{R}(r+u)]}{R^2(o)}-\frac{R(r+v)\delta[\hat{R}(o)]\delta[R(r)]}{R^3(o)}\right.$$

$$\left.-\frac{R(r)\delta[\hat{R}(o)]\delta[\hat{R}(r+v)]}{R^3(o)}+\frac{R(r)R(r+v)(\delta[\hat{R}(o)])^2}{R^4(o)}\right\}$$

$$=\frac{1}{\sigma^4}\{\mathrm{Cov}[\hat{R}(r),\hat{R}(r+v)]-\rho(r+v)\mathrm{Cov}[\hat{R}(o),\hat{R}(r)]$$

$$-\rho(r)\mathrm{Cov}[\hat{R}(o),\hat{R}(r+v)]+\rho(r)\rho(r+v)\mathrm{Var}[\hat{R}(o)]$$

利用式(3.1.16)近似关系得

$$\mathrm{Cov}[\hat{\rho}(r),\hat{\rho}(r+v)]\cong\frac{1}{N}\sum_{m=-\infty}^{\infty}\{\rho(m)\rho(m+v)+\rho(m+r+v)\rho(m-r)+2\rho(r)\rho(r+v)\rho^2(m)-2\rho(r)\rho(m)\rho(m-r-v)-2\rho(r+v)\rho(m)\rho(m-r)\} \tag{3-1-23}$$

当$v=0$时

$$\mathrm{Var}[\hat{\rho}(r)]\cong\frac{1}{N}\sum_{m=-\infty}^{\infty}\{\rho^2(m)+\rho(m+r)\rho(m-r)+2\rho^2(r)\rho^2(m)-4\rho(r)\rho(m)\rho(m-r)\} \tag{3-1-24}$$

巴特利还给出$\{x_t\}$为正态情况下的一般关系。以上结果表明，相关函数相邻点之间有很大的相关性，即便$|r|\to\infty$，而$\rho(r)\to 0$，但其估计$\hat{\rho}(r)$的相关性并不象$\rho(r)$那样快地衰减（对于$\hat{R}(r)$也有同样的结论）。实际上，若存在足够大的r，使得当$|s|\geqslant r$时$\rho(s)\cong 0$，则有

$$\mathrm{Cov}[\hat{\rho}(r),\hat{\rho}(r+v)]\cong\frac{1}{N}\sum_{m=-\infty}^{\infty}\rho(m)\rho(m+v) \tag{3-1-25}$$

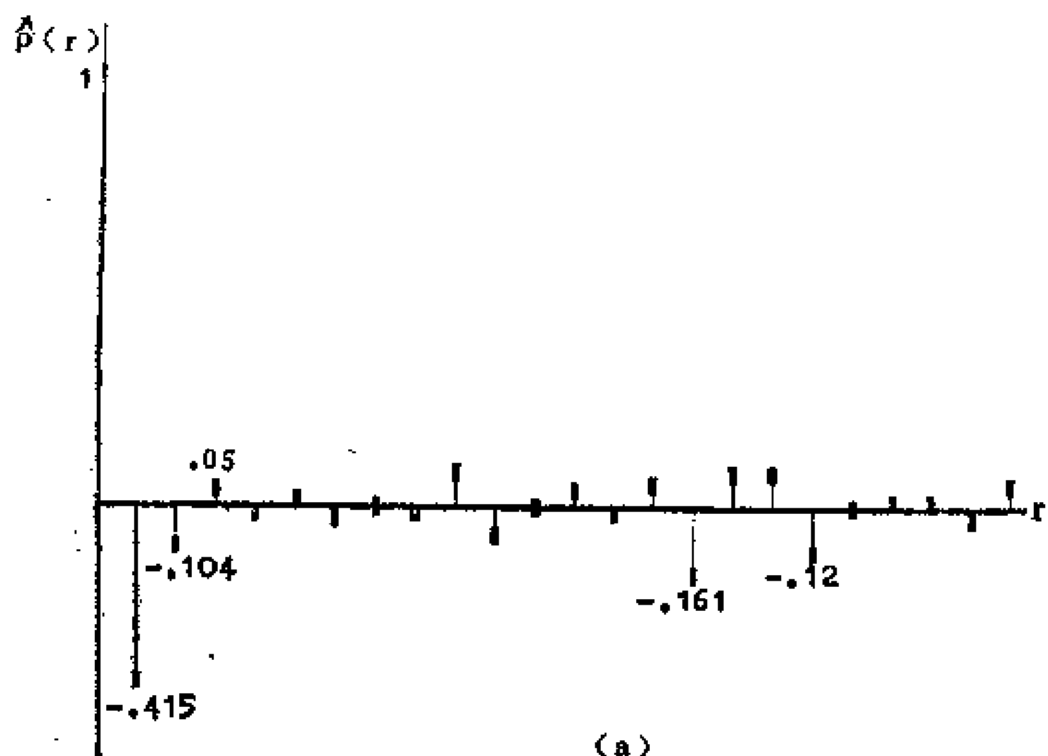
$$\mathrm{Var}[\hat{\rho}(r)]\cong\frac{1}{N}\sum_{m=-\infty}^{\infty}\rho^2(m) \tag{3-1-26}$$

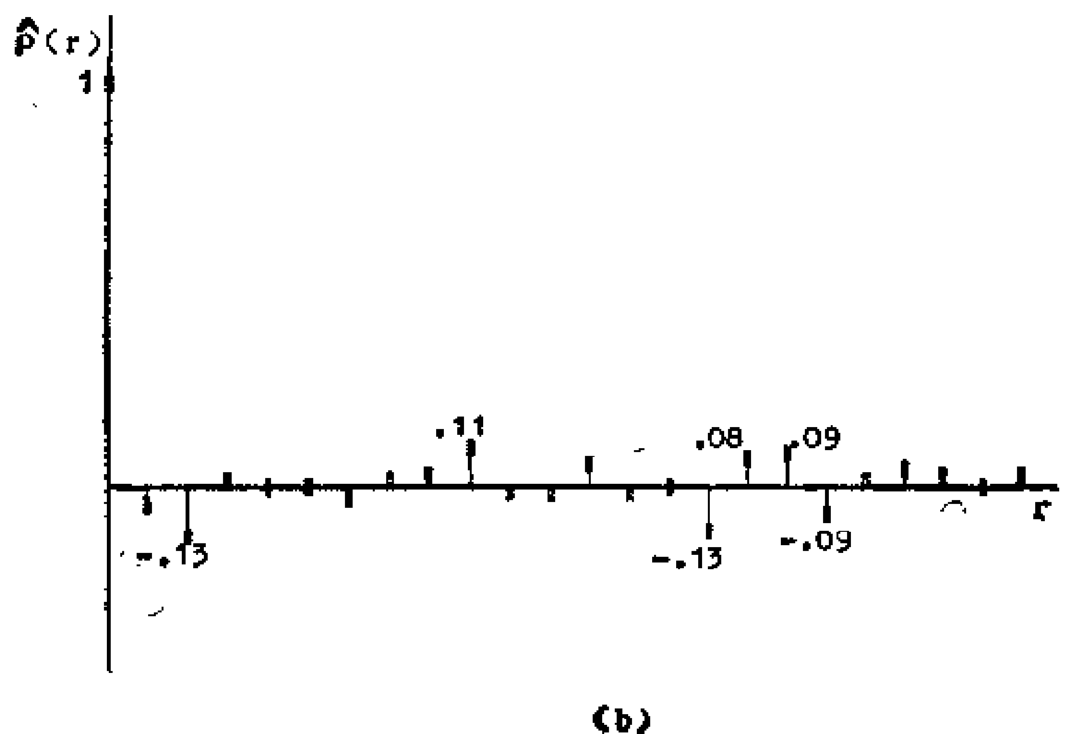
巴特利公式常常用来计算$\hat{\rho}(r)$的方差，实际上常取

$$\mathrm{Var}[\hat{\rho}(r)]\cong\frac{1}{N}\sum_{m=-K}^{K}\rho^2(m)=\frac{1}{N}\left(1+2\sum_{m=1}^{K}\rho^2(m)\right) \tag{3-1-27}$$

其中k可取为20。

图3—1a是由256个x_t的仿真值计算得出的样本相关函数$\hat{\rho}(r)(r=0,1,2,\cdots,12)$，$x_t$的仿真是对模型$x_t=\varepsilon_t-0.6\varepsilon_{t-1}$调用程序*DAGENT*得到的，其中$\varepsilon_t$是独立标准化正态变量（零均





(b)

图3-1　两个256点序列的样本自相关

值，单位方差)样本相关函数$\rho(r)$的计算调用了程序CORELA。

要判断$r>1$时ρ是否为零。根据式(3-1-26)

$$\begin{aligned}\text{Var}[\rho(r)] &\cong (1/256)[1+2\rho^2(1)]\\ &=[1+2(-0.415)^2]/256=0.00552\end{aligned}$$

故对于$r>1$

$$2\text{S.E.}[\rho(r)]\cong 2\sqrt{0.00525}=0.145$$

从图可知，$\rho(2)$，$\rho(3)$，…均不显著（$\rho(15)$略大），因此可以认为$r>1$时$\rho(r)=0$。

当过程为零均值白噪声时，式(3-1-26)变成(当$r>0$时)

$$\text{Var}[\rho(r)]\cong\frac{1}{N}$$

或

$$S.E.[\rho(r)]\cong\frac{1}{\sqrt{N}}$$

图3-1b为一标准化高斯白噪声的256个仿真值（仍调用程序DAGENT即可得到）的样本自相关图形

$$2\text{S.}E.[\rho(r)]=2/\sqrt{N}=0.13$$

3.2 模型参数的相关矩估计

矩估计在时间序列分析和建模中是一种常用的初步估计方法，尽管它不如我们后面将要介绍的最小二乘估计和极大似然估计精确和有效，但在有些情况下，这个初步估计（粗估计）和精估计很接近。

矩估计的基本思想是：假设随机变量x的概率分布含有k个未知数θ_1，θ_2，…，θ_K，我们把分布的前k阶矩表成θ_1，…，θ_k的函数，即

$$m_1=f_1(\theta_1,\cdots,\theta_k)$$

$$\vdots \qquad (3\text{-}2\text{-}1)$$

$$m_k = f_k(\theta_1, \cdots, \theta_k)$$

原则上可以由这些方程反过来将θ_1，…，θ_k表示成m_1，…，m_k的函数，即

$$\begin{aligned} \theta_1 &= g_1(m_1, \cdots, m_k) \\ &\vdots \\ \theta_k &= g_k(m_1, \cdots, m_k) \end{aligned} \qquad (3\text{-}2\text{-}2)$$

根据x的N个观察值，计算k个样本矩$\hat{m}_1, \cdots, \hat{m}_k$，并认为它是$m_1$，…，$m_k$的合理的估计。于是将$\hat{m}_1, \cdots, \hat{m}_k$代入式(3-2-2)便得$\theta_1$，…，$\theta_k$的估计值

$$\begin{aligned} \hat{\theta}_1 &= g_1(\hat{m}_1, \cdots, \hat{m}_k) \\ &\vdots \\ \hat{\theta}_k &= g_k(\hat{m}_1, \cdots, \hat{m}_k) \end{aligned} \qquad (3\text{-}2\text{-}3)$$

自协方差和自相关属于二阶矩，利用样本自协方差函数或样本自相关函数对x_t模型参数进行估计，就是一种矩估计。因为前文已经指出，R和ρ可通过$ARMA$过程的参数表达，把R或ρ以$\hat{R}$或$\hat{\rho}$代替，从中解出参数的估值$\hat{a}_1$，$\hat{a}_2$，…，$\hat{a}_n$，$\hat{b}_1$，$\hat{b}_2$，…，$\hat{b}_m$和$\hat{\sigma}_\varepsilon^2$就是$a_i$，$b_i$和$\sigma_\varepsilon^2$的矩估计。

1. AR模型参数的矩估计 利用$r=1$，…n的耶尔—瓦克尔方程式(2-3-50)(2-3-51)，并以$-\varphi_i$表参数a_i，当以$\hat{\rho}$代ρ可得参数估计

$$\begin{pmatrix} \hat{\varphi}_1 \\ \hat{\varphi}_2 \\ \vdots \\ \hat{\varphi}_n \end{pmatrix} = \begin{pmatrix} -\hat{a}_1 \\ -\hat{a}_2 \\ \vdots \\ -\hat{a}_n \end{pmatrix} = \begin{pmatrix} 1 & \hat{\rho}(1) \cdots \hat{\rho}(n-1) \\ \hat{\rho}(1) & 1 \quad \cdots \hat{\rho}(n-2) \\ \cdots\cdots & \cdots\cdots \\ \hat{\rho}(n-1)\hat{\rho}(n-2) & \cdots 1 \end{pmatrix}^{-1} \begin{pmatrix} \hat{\rho}(1) \\ \hat{\rho}(2) \\ \vdots \\ \hat{\rho}(n) \end{pmatrix} \qquad (3\text{-}2\text{-}4)$$

再将$\hat{a}_1, \cdots, \hat{a}_n$代替$a_1, \cdots, a_n$代入式(2-3-49)，得

$$\hat{\sigma}_\varepsilon^2=\hat{R}(0)+\sum_{i=1}^{n}\hat{a}_i\hat{R}(i)=\hat{R}(0)\left[1+\sum_{i=1}^{n}\hat{a}_i\hat{\rho}(i)\right]$$

例如AR(1)模型参数的矩估计为：

由$n=1$，耶尔一瓦克尔方程为

$$\hat{R}(1)+a_1\hat{R}(0)=0$$

故有

$$\hat{a}_1=-\hat{R}(1)/\hat{R}(0)=-\hat{\rho}(1)$$

而
$$\hat{\sigma}_\varepsilon^2=\hat{\sigma}_x^2(1-\hat{\rho}^2(1))$$

实际上，在讨论$AR(1)$模型时已有$\rho(r)=(-a_1)^{|r|}$故$\rho(1)=-a_1$。还可以看出

$$E(\hat{a}_1)=-E[\hat{\rho}(1)]=-\rho(1)=a_1$$

即$\hat{a}_1$是a_1的无偏估计。

AR(2)模型的矩估计如下

由$n=2$，耶尔一瓦克尔方程为

$$\hat{R}(1)+\hat{a}_1\hat{R}(0)+\hat{a}_2\hat{R}(1)=0$$
$$\hat{R}(2)+\hat{a}_1\hat{R}(1)+\hat{a}_2\hat{R}(0)=0$$

由此可得

$$\hat{a}_1=\hat{\rho}(1)(\hat{\rho}(2)-1)/(1-\hat{\rho}^2(1))$$
$$\hat{a}_2=(\hat{\rho}^2(1)-\hat{\rho}(2))\ /(1-\hat{\rho}^2(1))$$

而

$$\hat{\sigma}_\varepsilon^2\cong\hat{R}(0)(1+\hat{a}_1\hat{\rho}(1)+\hat{a}_2\hat{\rho}(2))$$

2. MA模型参数的矩估计 利用式(2-3-61)有(设$b_0=1$)

$$\hat{R}(r)=\begin{cases}\hat{\sigma}_\varepsilon^2(1+\hat{b}_1^2+\hat{b}_2^2+\cdots+\hat{b}_m^2) & r=0\\ \hat{\sigma}_\varepsilon^2(\hat{b}_r+\hat{b}_1\hat{b}_{r+1}+\cdots+\hat{b}_{m-r}\hat{b}_m) & r=1,2,\cdots,m\end{cases}\qquad(3\text{-}2\text{-}6)$$

由于这$m+1$个方程对参数而言是非线性的（二次方程），求解$\hat{b}_1,\cdots,\hat{b}_m$，$\hat{\sigma}_\varepsilon^2$的方法有以下三种：

(1)直接解法。$m=1$时有

$$\hat{\sigma}_\varepsilon^2(1+\hat{b}_1^2)=\hat{R}(0)$$

$$\hat{\sigma}_{\varepsilon}^{2}\hat{b}_{1}=\hat{R}(1)$$

可得$\hat{b}_1=\hat{R}(1)/\hat{\sigma}_{\varepsilon}^{2}$，及

$$\hat{\sigma}_{\varepsilon}^{4}-\hat{R}(0)\hat{\sigma}_{\varepsilon}^{2}+\hat{R}^{2}(1)=0$$

从而有

$$\hat{b}_1=2\hat{\rho}(1)/(1\pm\sqrt{1-4(\hat{\rho}(1))^2})$$
$$\hat{\sigma}_{\varepsilon}^{2}=\hat{R}(0)(1\pm\sqrt{1-4(\hat{\rho}(1)^2)})/2$$

可逆性条件($|\hat{b}_1|<1$)可以排除$\hat{b}_1$的多值性。注意到

$$\frac{2\hat{\rho}(1)}{1+\sqrt{1-4\hat{\rho}^2(1)}}\cdot\frac{2\hat{\rho}(1)}{1-\sqrt{1-4\hat{\rho}^2(1)}}=1$$

$$\left|\frac{2\hat{\rho}(1)}{1+\sqrt{1-4\hat{\rho}^2(1)}}\right|<1$$

故取

$$\hat{b}_1=\frac{2\hat{\rho}(1)}{1+\sqrt{1-4\hat{\rho}^2(1)}}\tag{3-2-7}$$

从而

$$\hat{\sigma}_{\varepsilon}^{2}=\hat{R}(0)(1+\sqrt{1-4\hat{\rho}^2(1)})/2$$

$m=2$时有

$$\hat{\sigma}_{\varepsilon}^{2}(1+\hat{b}_1^2+\hat{b}_2^2)=\hat{R}(0)$$
$$\hat{\sigma}_{\varepsilon}^{2}(\hat{b}_1+\hat{b}_1\hat{b}_2)=\hat{R}(1)$$
$$\hat{\sigma}_{\varepsilon}^{2}\hat{b}_2=\hat{R}(2)$$

可得

$$\hat{b}_2=\hat{R}(2)/\hat{\sigma}_{\varepsilon}^{2},\qquad \hat{b}_1=\hat{R}(1)/(\hat{\sigma}_{\varepsilon}^{2}+\hat{R}(0))$$
$$\hat{\sigma}_{\varepsilon}^{2}\{1+\hat{R}^2(2)/\hat{\sigma}_{\varepsilon}^{4}+\hat{R}^2(1)/(\hat{\sigma}_{\varepsilon}^{2}+\hat{R}(2))^2\}=\hat{R}(0)$$

这是$\hat{\sigma}_{\varepsilon}^{2}$的四次方程，有四个根，因此$\hat{b}_1$，$\hat{b}_2$有四种可能的解。利用可逆性条件，经过烦琐的讨论可以得出唯一解

$$\hat{b}_2=-\frac{1}{2}+\frac{1}{4\hat{\rho}(2)}+\frac{1}{2\hat{\rho}(2)}\sqrt{(\hat{\rho}(2)+\frac{1}{2})^2-\hat{\rho}^2(1)}$$

$$\pm\sqrt{\left(\frac{1}{2}-\frac{1}{4\rho(2)}-\frac{1}{2\rho(2)}\sqrt{(\rho(2)+\frac{1}{2})^2-\rho^2(1)^2-1}\right.}$$

(3-2-9)

式中"±"号依$\rho(2)>0$或$\rho(2)<0$而分别取一或+

$$\hat{b}_1=\frac{\rho(1)\hat{b}_2}{\rho(2)(1+\hat{b}_2)},\qquad \sigma_\varepsilon^2=\hat{R}(0)\rho(2)/\hat{b}_2 \qquad (3\text{-}2\text{-}10)$$

由于按直接法计算中要解$2m$次方程，在$m\geqslant3$以后一般只能用数值解法。

(2)线性迭代法。将式(3-2-6)写成

$$\sigma_\varepsilon^2=\hat{R}(0)/(1+\hat{b}_1^2+\cdots+\hat{b}_m^2) \qquad (3\text{-}2\text{-}11)$$

$$\hat{b}_r=\hat{R}(r)/\sigma_\varepsilon^2-\hat{b}_1\hat{b}_{r+1}-\cdots-\hat{b}_{m-r}\hat{b}_m \qquad (3\text{-}2\text{-}12)$$

$$r=1,2,\cdots,m$$

给定$\hat{b}_1$，$\cdots\hat{b}_m$和σ_ε^2的一组初值（如$\hat{b}_1=\cdots=\hat{b}_m=0$，$\sigma_\varepsilon^2=R(0)$或$\hat{b}_1=1$，$\hat{b}_2=\cdots=\hat{b}_m=0$，$\sigma_\varepsilon^2=\hat{R}(0)/2$等）。代入式(3-2-11)，(3-2-12)右边，左边所得的值为一步迭代值$\sigma_\varepsilon^{2(1)}$，$b_1^{(1)},\cdots,b_m^{(1)}$。再将它们代入上式右边，得第二步迭代值$\sigma_\varepsilon^{2(2)}$，$\hat{b}_1^{(2)},\cdots\hat{b}_m^{(2)}$，依此类推，直到相邻两次迭代值结果相差不大时停止迭代，取最后结果作为上式近似解。

(3) 牛顿—拉夫森（*Newton-Raphson*）算法[4]

3. ARMA模型参数的矩估计 先看一种简单情况，即ARMA(1,1)过程

$$x_t+ax_{t-1}=\varepsilon_t+b\varepsilon_{t-1}$$

其平稳和可逆条件是$|a|<1$，$|b|<1$。由2.3.6节例2

$$\sigma_x^2=R(0)=\frac{1-2ab+b^2}{1-a^2}\sigma_\varepsilon^2$$

$$\sigma_x^2\rho(1)=R(1)=\frac{(b-a)(1-ab)}{1-a^2}\sigma_\varepsilon^2$$

$$\sigma_x^2\rho(r)=R(r)=-aR(r-1) \qquad r>1$$

利用估计的相关函数可得

$$\hat{a}=-\hat{R}(2)/\hat{R}(1)=-\hat{\rho}(2)/\hat{\rho}(1) \qquad (3\text{-}2\text{-}13)$$

将$R(1)$和$R(0)$两式相除可得

$$\rho(1)=\frac{(b-a)(1-ab)}{1-2ab+b^2}$$

或

$$b^2-\left\{\frac{2\rho(1)a+a^2+1}{\rho(1)+a}\right\}b+1=0$$

故

$$\hat{b}=\frac{1}{2}(c\pm\sqrt{c^2-4}) \qquad (3\text{-}2\text{-}14)$$

其中

$$c=(\hat{a}^2-2\hat{\rho}(2)+1)/(\hat{\rho}(1)+\hat{a})$$

$$\hat{\sigma}_\varepsilon^2=\hat{R}(0)(1-\hat{a}^2)/(1-2\hat{a}\hat{b}+\hat{b}^2) \qquad (3\text{-}2\text{-}15)$$

对于高阶的ARMA模型其参数的矩估计可以采用以下近似方法。

第一步：先给出AR部分a_1，$\cdots a_n$的矩估计

利用(2-3-72)式中$r>m$时相关函数方程组来求解$\hat{a}_1$，$\cdots$，$\hat{a}_n$，即

$$\begin{pmatrix}-\hat{a}_1\\-\hat{a}_2\\\vdots\\-\hat{a}_n\end{pmatrix}=\begin{pmatrix}\hat{\varphi}_1\\\hat{\varphi}_2\\\vdots\\\hat{\varphi}_n\end{pmatrix}$$

$$=\begin{pmatrix}\hat{\rho}(m) & \hat{\rho}(m-1)\cdots\cdots\hat{\rho}(m-n+1)\\ \hat{\rho}(m+1) & \hat{\rho}(m)\cdots\cdots\cdots\hat{\rho}(m-n+2)\\ \cdots\cdots\cdots & \\ \hat{\rho}(m+n-1) & \hat{\rho}(m+n-2)\cdots\hat{\rho}(m)\end{pmatrix}^{-1}\begin{pmatrix}\hat{\rho}(m+1)\\\hat{\rho}(m+2)\\\vdots\\\hat{\rho}(m+n)\end{pmatrix}$$

在这里，我们没有考虑$r\leqslant m$时的方程式，也就是暂不考虑滑动平均部分的作用，因此解出的$\hat{a}_i$是一种近似。

第二步：令$y_t = x_t + \hat{a}_1 x_{t-1} + \cdots + \hat{a}_n x_{t-n}$

其协方差函数为

$$R_y(r) = E[y_t y_{t+r}]$$
$$= E[(x_t + \hat{a}_1 x_{t-1} + \cdots + \hat{a}_n x_{t-n})(x_{t+r} + \hat{a}_1 x_{t+r-1} + \cdots + \hat{a}_n x_{t+r-n})]$$
$$= \sum_{i;\, j=0}^{n} \hat{a}_i \hat{a}_j R(r+j-i) \qquad \hat{a}_0 = 1$$

再将R以$\hat{R}$代替代入上式作为R_y的样本函数$\hat{R}_y$，即

$$\hat{R}_y(r) = \sum_{i;\, j=0}^{n} \hat{a}_i \hat{a}_j \hat{R}(r+j-i)$$

第三步：把y_t近似看做MA(m)过程，即

$$y_t \cong \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_m \varepsilon_{t-m}$$

利用前面关于MA参数的矩估计方法解下列方程

$$\hat{R}_y(0) = \hat{\sigma}_\varepsilon^2 (1 + \hat{b}_1^2 + \cdots + \hat{b}_m^2)$$
$$\hat{R}_y(r) = \hat{\sigma}_\varepsilon^2 (\hat{b}_r + \hat{b}_1 \hat{b}_{r+1} + \cdots + \hat{b}_{m-r} \hat{b}_m) \quad (3\text{-}2\text{-}16)$$
$$r = 1, 2, \cdots, m$$

在附录二中给出了按照这三个步骤计算ARMA模型参数的程序ARMAME。对AR模型及MA模型可设定$m=0$或$n=0$，由本程序进行参数的矩估计

3.3 模型参数的最小二乘估计(LS估计)

1. 最小二乘法 没有线性模型

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_n x_{in} + e_i \quad i = 1, \cdots, N \quad (3\text{-}3\text{-}1)$$

其中$y_1, \cdots, y_N$为观察数据，而自变量$x_{i1}, \cdots x_{in}$已知，$\beta_1, \cdots, \beta_n$是未知的待估计参数，e_i是不相关的零均值误差—模型(3-3-1)写成矩阵形式

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} x_{11}x_{12}\cdots x_{1n} \\ x_{21}x_{22}\cdots x_{2n} \\ \cdots\cdots \\ x_{N1}x_{N2}\cdots x_{Nn} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_N \end{pmatrix} \qquad (3\text{-}3\text{-}2)a$$

或

$$Y = X\boldsymbol{\beta} + e$$

能使残差平方和

$$Q(\boldsymbol{\beta}) = Q(\beta_1, \beta_2, \cdots, \beta_n) \equiv \sum_{i=1}^{N} (y_i - \beta_1 x_{i1} - \beta_2 x_{i2} - \cdots - \beta_n x_{in})^2 \qquad (3\text{-}3\text{-}3)$$

$$= \sum_{i=1} e_i^2$$

达到极小的$\widehat{\boldsymbol{\beta}} = (\beta_1, \cdots, \beta_n)^T$称为参数$\widehat{\boldsymbol{\beta}}$的最小二乘估计。由

$$\begin{aligned} Q(\boldsymbol{\beta}) &= e^T e = (Y - X\boldsymbol{\beta})^T (Y - X\boldsymbol{\beta}) \\ &= [Y - X\widehat{\boldsymbol{\beta}} - X(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}})]^T [Y - X\widehat{\boldsymbol{\beta}} - X(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}})] \\ &= Q(\widehat{\boldsymbol{\beta}}) + (\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}})^T X^T X (\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}) + (X^T X \widehat{\boldsymbol{\beta}} - X^T Y)^T (\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}) \\ &\quad - (\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}})^T (X^T Y - X^T X \widehat{\boldsymbol{\beta}}) \end{aligned}$$

其中

$$Q(\widehat{\boldsymbol{\beta}}) = (Y - X\widehat{\boldsymbol{\beta}})^T (Y - X\widehat{\boldsymbol{\beta}})$$

表示参数取为$\widehat{\boldsymbol{\beta}}$时的残差平方和。可以看出当

$$(X^T X)\widehat{\boldsymbol{\beta}} = X^T Y \qquad (3\text{-}3\text{-}4)$$

时$Q(\boldsymbol{\beta})$达到最小。将$\boldsymbol{\beta}$的最小二乘估计记作$\widehat{\boldsymbol{\beta}}_{LS}$，则有

$$\widehat{\boldsymbol{\beta}}_{LS} = (X^T X)^{-1} X^T Y \qquad (3\text{-}3\text{-}5)$$

它的数学期望为

$$\begin{aligned} E[\widehat{\boldsymbol{\beta}}_{LS}] &= E[(X^T X)^{-1} X^T Y] = (X^T X)^{-1} X^T E[Y] \\ &= (X^T X)^{-1} X^T X \boldsymbol{\beta} = \boldsymbol{\beta} \end{aligned} \qquad (3\text{-}3\text{-}6)$$

这里利用了$E[e]=0$且e与X无关的条件。该式表明$\widehat{\boldsymbol{\beta}}_{LS}$是$\boldsymbol{\beta}$的无

偏估计。

下面导出$\widehat{\beta}_{LS}$的方差、协方差阵。在和上面提到的相同条件下，令

$$A=(X^TX)^{-1}X^T$$

则
$$\widehat{\beta}_{LS}=\beta+Ae \tag{3-3-7}$$

故
$$\begin{aligned}\mathrm{cov}[\widehat{\beta}_{LS}]&=E[(\widehat{\beta}_{LS}-E[\widehat{\beta}_{LS}])(\widehat{\beta}_{LS}-E[\widehat{\beta}_{LS}])^T]\\&=E[(\widehat{\beta}_{LS}-\beta)(\widehat{\beta}_{LS}-\beta)^T]=E[(Ae)(Ae)^T]\\&=AE[ee^T]A^T\end{aligned} \tag{3-3-8}$$

如果误差e是无关的，且其方差阵为常值，即 $E[ee^T]=\sigma_e^2\mathrm{I}$，则有

$$\mathrm{Cov}[\widehat{\beta}_{LS}]=\sigma_e^2(X^TX)^{-1} \tag{3-3-9}$$

这也是最小二乘意义下的最小协方差阵。由式(3-3-9)的主对角线元素（即$\widehat{\beta}_{LS}$的方差)可以计算β_{LS}的置信区间。

将$\widehat{\beta}=\widehat{\beta}_{LS}$时的$Q(\widehat{\beta})$记作$Q_{LS}$,表示LS估计时的最小残差平方和的估计，则

$$\begin{aligned}Q_{LS}&=(Y-X\widehat{\beta}_{LS})^T(Y-X\widehat{\beta}_{LS})\\&=Y^TY-\widehat{\beta}_{LS}^TX^TY-Y^TX\widehat{\beta}_{LS}+\widehat{\beta}_{LS}^TX^TX\widehat{\beta}_{LS}\\&=Y^TY-\widehat{\beta}_{LS}^TX^TX\widehat{\beta}_{LS}=Y^TY-\widehat{\beta}_{LS}^TX^TY\end{aligned} \tag{3-3-10}$$

实际计算时常用式(3-3-10)最后一个等式，因为X^TY在求$\widehat{\beta}_{LS}$时已经获得。

如果e为正态分布，并考虑到计算$\widehat{\beta}_{LS}$时利用了n个线性关系式(3-3-4)，则Q_{LS}是自由度为$N-n$的χ^2变量，故残差方差σ_e^2的无偏估计为

$$\hat{\sigma}_e^2=\frac{Q_{LS}}{N-n} \tag{3-3-11}$$

以上讨论的是线性模型(对于β)，式(3-3-4)也称为正规方程，它也可通过将Q对β_1，…，β_n取偏导数，并令其为零而获得。

如果考虑一般情况，即

$$y_i=f_i(\boldsymbol{\beta};\ y_1,\ \cdots,\ y_{i-1})+e_i \quad i=1,\ \cdots,\ N \qquad (3\text{-}3\text{-}12)$$

其中$\boldsymbol{\beta}=(\beta_1,\ \beta_2,\ \cdots,\ \beta_n)^T$为未知参数矢量，$f_i$是关于$\boldsymbol{\beta}$的非线性函数。根据观察值$y_1 y_2,\ \cdots,\ y_N$使残差平方和

$$Q(\boldsymbol{\beta})=\sum_{i=1}^{N}[y_i-f_i(\boldsymbol{\beta};\ y_1,\ \cdots,\ y_{i-1})]^2 \qquad (3\text{-}3\text{-}13)$$

达到最小的解$\widehat{\boldsymbol{\beta}}$称为$\beta$的最小二乘估计。这时$\beta$的求解比线性要复杂得多。我们将在ARMA参数估计中具体讨论。

2．AR模型参数的LS估计 将$AR(n)$模型式(2-3-43)改写成（设$\mu=0$）

$$x_t=\varphi_1 x_{t-1}+\cdots+\varphi_n x_{t-n}+\varepsilon_t \qquad (3\text{-}3\text{-}14)$$

今已知$x_1,\ x_2,\ \cdots,\ x_N$，求$\varphi_1,\ \varphi_2,\ \cdots,\ \varphi_n$的LS估计。由于$t<1$时的$x_t$是未知的，式(3-3-14)实际所包含的方程是

$$\begin{aligned}
x_{n+1}&=\varphi_1 x_n+\varphi_2 x_{n-1}+\cdots+\varphi_n x_1+\varepsilon_{n+1}\\
x_{n+2}&=\varphi_1 x_{n+1}+\varphi_2 x_{n+2}+\cdots+\varphi_n x_2+\varepsilon_{n+2}\\
&\vdots\\
x_N&=\varphi_1 x_{N-1}+\varphi_2 x_{N-2}+\cdots+\varphi_n x_{N-n}+\varepsilon_N
\end{aligned} \qquad (3\text{-}3\text{-}15)$$

亦即认为$\varepsilon_1=\varepsilon_2=\cdots=\varepsilon_n=0$、当$n\ll N$时，忽略前$t$个$\varepsilon_t$对整个平方和$\sum_{t=1}^{N}\varepsilon_t^2$影响很小

令

$$Y=[x_{n+1},\ x_{n+2},\ \cdots,\ x_N]^T$$

$$\boldsymbol{\beta}=[\varphi_1,\ \varphi_2,\ \cdots,\ \varphi_n]^T$$

$$X=\begin{pmatrix} x_n & x_{n-1} & \cdots & x_1\\ x_{n+1} & x_n & \cdots & x_2\\ \cdots & & & \\ x_{N-1} & x_{N-2} & \cdots & x_{N-n}\end{pmatrix} \qquad (3\text{-}3\text{-}16)$$

$$\mathbf{e}=[\varepsilon_{n+1},\ \varepsilon_{n+2},\ \cdots,\ \varepsilon_N]^T$$

可见式(3-3-15)和多元($n\geqslant 1$)回归模型具有相同的形式。在回归分析中，自变量X是确定的，Y和e是随机变量，在AR模型式(3-3-14)中，x_{t-1}，x_{t-2}，…，自然也是随机变量，但在$t-1$时刻它们均已确定不变，所以$AR(n)$模型可以看作是条件线性回归模型，仍可用多元回归分析中的有关方法进行参数估计。

利用式(3-3-5)便可得出AR模型参数估计的具体表达式为

$$\begin{pmatrix}-\hat{a}_1\\-\hat{a}_2\\\vdots\\-\hat{a}_n\end{pmatrix}=\begin{pmatrix}\hat{\varphi}_1\\\hat{\varphi}_2\\\vdots\\\hat{\varphi}_n\end{pmatrix}=\begin{pmatrix}\sum\limits_{t=n}^{N-1}x_t^2 & \sum\limits_{t=n}^{N-1}x_t x_{t-1}\cdots & \sum\limits_{t=n}^{N-1}x_t x_{t-(n-1)}\\ \sum\limits_{t=n}^{N-1}x_{t-1}x_t & \sum\limits_{t=n}^{N-1}x_{t-1}^2\cdots & \sum\limits_{t=n}^{N-1}x_{t-1}x_{t-(n-1)}\\ \vdots & \vdots & \vdots\\ \sum\limits_{t=n}^{N-1}x_{t-(n-1)}x_t & \sum\limits_{t=n}^{N-1}x_{t-(n-1)}x_{t-1}\cdots & \sum\limits_{t=n}^{N-1}x_{t-(n-1)}^2\end{pmatrix}^{-1}\begin{pmatrix}\sum\limits_{t=n}^{N-1}x_t x_{t+1}\\ \sum\limits_{t=n}^{N-1}x_{t-1}x_{t+1}\\ \vdots\\ \sum\limits_{t=n}^{N-1}x_{t-(n-1)}x_{t+1}\end{pmatrix} \tag{3-3-17}$$

将上式$\sum$的上下限改为$\sum\limits_{t=1}^{N-n}$并相应改变x的下标，并令

$$\hat{R}'(q-p)=\frac{1}{N}\sum_{t=1}^{N-n}x_{t+n-p}x_{t+n-q} \tag{3-3-18}$$

则式(3-3-17)可改写成如下近似式

$$\begin{pmatrix} -\hat{a}_1 \\ -\hat{a}_2 \\ \vdots \\ -\hat{a}_n \end{pmatrix} = \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_n \end{pmatrix}$$

$$\cong \begin{pmatrix} \hat{R}'(0) & \hat{R}'(1) \cdots \hat{R}'(n-1) \\ \hat{R}'(1) & \hat{R}'(0) \cdots \hat{R}'(n-2) \\ \cdots\cdots & \\ \hat{R}'(n-1) & \hat{R}'(n-2) \cdots \hat{R}'(0) \end{pmatrix}^{-1} \begin{pmatrix} \hat{R}'(1) \\ \hat{R}'(2) \\ \vdots \\ \hat{R}'(n) \end{pmatrix} \tag{3-3-19}$$

不难看出，当N很大时式(3-3-18)和由式(3-1-5)所定义的$\hat{R}$是近乎相等的(即$\hat{R}'\cong\hat{R}$)，因此LS估计和矩估计在样本量充分大时是十分接近的，而后者的计算量要小得多。

上面讨论中都是从$t=n+1$开始取观察值的，即在N个观察值中只取$N-n$个有效观察值，相应地

$$Q_{LS}=\sum_{t=n+1}^{N}\varepsilon_t^2 \tag{3-3-20}$$

因此残差方差的无偏估计为

$$\hat{\sigma}_\varepsilon^2=\frac{Q_{LS}}{N-2n} \tag{3-3-21}$$

此外，以上还假定x_t的均值为零。如果$E[x_t]=\mu\neq0$，则AR模型可写成

$$x_t-\mu+a_1(x_{t-1}-\mu)+\cdots+a_n(x_{t-n}-\mu)=\varepsilon_t \tag{3-3-22}$$

残差平方和为

$$Q(\mu,a_1,\cdots a_n)=\sum_{t=n+1}^{N}\varepsilon_t^2=\sum_{t=n+1}^{N}[x_t-\mu+a_1(x_{t-1}-\mu)+\cdots+a_n(x_{t-n}-\mu)]^2 \tag{3-3-23}$$

将Q对μ，a_1，…，a_n求极小，可得$n+1$个方程式

$$\sum_{t=n+1}^{N}[x_t-\hat{\mu}+\hat{a}_1(x_{t-1}-\hat{\mu})+\cdots+\hat{a}_n(x_{t-n}-\hat{\mu})]=0 \quad (3\text{-}3\text{-}24)$$

和

$$\sum_{t=n+1}^{N}[x_t-\hat{\mu}+\hat{a}_1(x_{t-1}-\hat{\mu})+\cdots+\hat{a}_n(x_{t-n}-\hat{\mu})](x_{t-i}-\hat{\mu})=0$$

$$i=1,\ \cdots,\ n \qquad (3\text{-}3\text{-}25)$$

由式(3-3-24)可得均值的估计为

$$\hat{\mu}=\frac{\bar{x}_1+\hat{a}_1\bar{x}_2+\hat{a}_2\bar{x}_3+\cdots+\hat{a}_n\bar{x}_{n+1}}{1+\hat{a}_1+\cdots+\hat{a}_n} \qquad (3\text{-}3\text{-}26)$$

其中

$$\bar{x}_{i+1}=\frac{1}{N-n}\sum_{t=n-i+1}^{N-i}x_t \qquad i=0,1,\cdots,n$$

$\bar{x}_1$，$\bar{x}_2$，…，$\bar{x}_{n+1}$分别代表不同组（每组$N-n$个）观察值的平均，当$N\gg n$时，它们都和全部平均$\bar{x}=\frac{1}{N}\sum_{t=1}^{N}x_t$很接近，即$\mu$的近似估计为

$$\hat{\mu}=\bar{x} \qquad (3\text{-}3\text{-}27)$$

于是n个方程式(3-3-25)近似为

$$\sum_{t=n+1}^{N}[x_t-\bar{x}+\hat{a}_1(x_{t-1}-\bar{x})+\cdots+\hat{a}_n(x_{t-n}-\bar{x})](x_{t-i}-\bar{x})=0$$

$$i=1,2,\cdots,n \qquad (3\text{-}3\text{-}28)$$

从该式求解$\hat{a}_1$，…，$\hat{a}_n$的公式和式(3-3-17)一样，只要将其中的x_t减去$\bar{x}$即可。

将μ和a_i的估值替代式(3-3-23)中相应的值即得Q_{LS}，这时的自由度为$N-2n-1$，故

$$\hat{\sigma}_{\varepsilon}^{2}=\frac{Q_{LS}}{N-2n-1}=\frac{1}{N-2n-1}Q(\hat{\mu},\hat{a}_1,\cdots,\hat{a}_n) \tag{3-3-29}$$

如果取$x_0=x_{-1}=\cdots=x_{-n+1}=0$，而$\varepsilon_t$的计算为：

$\varepsilon_1=x_1$；$\varepsilon_2=x_2+a_1x_1$；$\cdots,\varepsilon_N=x_N+a_1x_{N-1}+\cdots+a_nx_{N-n}$

这时t由1开始共取N个观察值，$Q=\sum_{t=1}^{N}\varepsilon_t^2$，则相应地有

$$\hat{\sigma}_{\varepsilon}^{2}=\frac{1}{N-n-1}Q(\hat{\mu},\ \hat{a}_1,\cdots,\ \hat{a}_n) \tag{3-3-30}$$

和

$$\hat{\sigma}_{\varepsilon}^{2}=\frac{1}{N-n}Q(\hat{a}_1,\ \cdots,\ \hat{a}_n) \tag{3-3-31}$$

3. ARMA模型参数的LS估计 本节讨论包括滑动平均模型在内的$ARMA$模型参数的最小二乘估计，这种模型可写为

$$x_t=-a_1x_{t-1}-a_2x_{t-2}-\cdots-a_nx_{t-n}+b_1\varepsilon_{t-1}+\cdots+b_m\varepsilon_{t-m}+\varepsilon_t \tag{3-3-32}$$

或

$$\varepsilon_t=x_t+a_1x_{t-1}+\cdots+a_nx_{t-n}-b_1\varepsilon_{t-1}-\cdots-b_m\varepsilon_{t-m} \tag{3-3-33}$$

利用逆转关系式(2-3-78)可以将$\varepsilon_{t-i}(i=1,\ \cdots,\ m)$化成$x_{t-i}$，$x_{t-i-1}\cdots$的线性组合，但是由于逆函数本身是$a_1$，$\cdots$，$a_n$，$b_1$，$\cdots$，$b_m$的函数，因此最终得到的方程式对于参数$(a_i，b_i)$来说必然是非线性的。这一点与$AR$模型的方程式〔见式(3-3-15)或(3-3-22)〕是截然不同的。

在需要估计均值时，$ARMA$模型表达式写成

$$x_t-\mu=-\sum_{i=1}^{n}a_i(x_{t-i}-\mu)+\sum_{i=1}^{m}b_i\varepsilon_{t-i}+\varepsilon_t \tag{3-3-34}$$

总之，在一般情况下有

$$x_t = f_t(\mu, a_i, b_i, x_i) + \varepsilon_t \tag{3-3-35}$$

或

$$\varepsilon_t = x_t - f_t(\mu, a_i, b_i, x_i) \tag{3-3-36}$$

其中$t=1, 2, \cdots, N$或$t=n+1, n+2, \cdots, N$，这里f_t是对于参数的非线性函数关系。若根据$N-n$个实际观察值可以将式(3-3-36)表示成

$$Y = F(X, \beta) + e \tag{3-3-37}$$

其中

$Y = [x_{n+1}, x_{n+2}, \cdots, x_N]^T$，$F = [f_{n+1}, f_{n+2}, \cdots, f_N]^T$，

$e = [\varepsilon_{n+1}, \varepsilon_{n+2}, \cdots, \varepsilon_N]^T$

X表示观察值$x_1, \cdots, x_N$，β表示参数$\mu, a_1, \cdots, a_n, b_1, \cdots, b_m$（当取$N$个观察时，$Y$，$F$，$e$诸元素的下标从1开始到$N$）

残差平方和为

$$Q = [Y - F(X, \beta)]^T [Y - F(X, \beta)] \tag{3-3-38}$$

由于F不是β的线性函数，所以根据使Q达极小来求β，便遇到了求解非线性最小二乘的问题。

AR模型可以看成是一种特例，它有$F(X, \beta) = X\beta$，下面将对非线性参数模型采用某种线性化的方法，就是要使$F(X, \beta)$近似地表示成参数向量β和另一个与β无关的量的乘积。基本的线性化方法，就是将$F(X, \beta)$在$\beta = \beta_0$处作台劳展开，然后取它的线性部分，这就是高斯—牛顿法。

由于在计算中X为已知量，下面把$F(X, \beta)$写成$F(\beta)$，以突出它是β的函数。其展开式的线性近似为

$$F(\beta) = \begin{pmatrix} f_{n+1}(\beta_1, \beta_2, \cdots, \beta_q) \\ f_{n+2}(\beta_1, \beta_2, \cdots, \beta_q) \\ \vdots \\ f_N(\beta_1, \beta_2, \cdots, \beta_q) \end{pmatrix}$$

$$
=\begin{pmatrix} f_{n+1}(\beta_{10},\beta_{20},\cdots\beta_{q0}) \\ f_{n+2}(\beta_{10},\beta_{20},\cdots\beta_{q0}) \\ \vdots \\ f_N(\beta_{10},\beta_{20},\cdots,\beta_{q0}) \end{pmatrix}+\begin{pmatrix} \frac{\partial f_{n+1}}{\partial \beta_1}, & \frac{\partial f_{n+1}}{\partial \beta_2}, & \cdots\frac{\partial f_{n+1}}{\partial \beta_q} \\ \frac{\partial f_{n+2}}{\partial \beta_1}, & \frac{\partial f_{n+2}}{\partial \beta_2}, & \cdots\frac{\partial f_{n+2}}{\partial \beta_q} \\ \vdots & & \\ \frac{\partial f_N}{\partial \beta_1}, & \frac{\partial f_N}{\partial \beta_1}, & \cdots\frac{\partial f_N}{\partial \beta_q} \end{pmatrix}_{\beta=\beta_0}\begin{pmatrix} \beta_1-\beta_{10} \\ \beta_2-\beta_{20} \\ \vdots \\ \beta_q-\beta_{q0} \end{pmatrix} \tag{3-3-39}
$$

其中$\beta=[\beta_1,\beta_2,\ \cdots,\ \beta_q]^T$，$\beta_0=[\beta_{10},\beta_{20},\cdots,\beta_{q0}]^T$，$\beta_1,\beta_2,\cdots,\beta_q$分别为$\mu,a_1,\ \cdots a_n,b_1,\ \cdots,b_m$诸参数。当估计均值时，$q=n+m+1$；不估计均值时，$q=n+m$。把式(3-3-39)右边导数阵称为灵敏度矩阵，并记作$X(\beta_0)$，则有

$$F(\beta)=F(\beta_0)+X(\beta_0)(\beta-\beta_0) \tag{3-3-40}$$

于是，在$\beta=\beta_0$处残差平方和为

$$Q=[Y-F(\beta_0)-X(\beta_0)(\beta-\beta_0)]^T[Y-F(\beta_0)-X(\beta_0)(\beta-\beta_0)] \tag{3-3-41}$$

由于在包括β_0的区域附近，F可以近似地看作是线性的，那么能使Q更小的参数估值$\hat{\beta}$应满足

$$\hat{\beta}-\beta_0=[X^T(\beta_0)X(\beta_0)]^{-1}X^T(\beta_0)[Y-F(\beta_0)] \tag{3-3-42}$$

或

$$\hat{\beta}=\beta_0+[X^T(\beta_0)X(\beta_0)]^{-1}X^T(\beta_0)[Y-F(\beta_0)] \tag{3-3-43}$$

实际计算中，$\hat{\beta}$是迭代式地进行，以β_0作为$\hat{\beta}$的初始估计值，用k和$k+1$表示第k次和第$k+1$次的计算结果，则有

$$\hat{\beta}^{(k+1)}=\hat{\beta}^{(k)}+P^{(k)}X^{T(k)}(Y-F^{(k)}) \tag{3-3-44}$$

其中

$$P^{(k)}=[X^{T(k)}X^{(k)}]^{-1} \tag{3-3-45}$$

通常规定最大迭代次数，或规定当$\widehat{\beta}$的每一分量都变化不大时停止迭代，并取最后的结果作为$\widehat{\beta}_{LS}$。参数的相对变化用下面的不等式表示：

$$\frac{|\beta_i^{(k+1)}-\beta_i^{(k)}|}{|\beta_i^{(k)}|+\delta_1}<\delta \quad i=1,2,\cdots,q \tag{3-3-46}$$

δ为设定值，例如可取其为10^{-5}或10^{-4}，为避免$\beta_i^{(k)}\approx 0$而产生的麻烦，我们引入δ_1(可取为10^{-100})。

式(3-3-42)常称为步长，$P=[X^T(\beta_0)X(\beta_0)]^{-1}$称为步长系数，而在$\beta_0$处Q的梯度为$-2X^T(\beta_0)[Y-F(\beta_0)]$。因此$\widehat{\beta}$是在$\beta_0$的基础上沿梯度的相反方向变化而得到的。高斯—牛顿法把平方和Q的曲面局部近似为二次曲面，当初始值β_0远离Q的真正最低点，且曲面形状较不规则，或当X的情况“不好”时，P值可能很大（求P要作求逆运算），从而使参数调整的步长加大，以致迭代过程剧烈“振荡”或发散。因此在初始估计很粗糙时往往要考虑用别的算法，只是在接近Q的真正最小值时才用高斯—牛顿法。

一种称为最速下降法的迭代算法是将步长系数定为常数λ^{-1}，即

$$\widehat{\beta}^{(k+1)}=\widehat{\beta}^{(k)}+\lambda^{-1}X^{(k)}(Y-F^{(k)}) \tag{3-3-47}$$

当λ很大时可以使步长较小。这里无需矩阵求逆的运算。然而，固定的步长系数并不总是理想的，为了使步长系数能受到控制，利温伯格(*Levenberg*)试图通过在式(3-3-43)中附加一项来计算步长，即

$$\widehat{\beta}^{(k+1)}=\widehat{\beta}^{(k)}+[X^{T(k)}X^{(k)}+\lambda^{(k)}D^{(k)}]^{-1}X^{T(k)}(Y-F^{(k)}) \tag{3-3-48}$$

特别是当λD的各主对角线元素比X^TX的相应元素都大的情况下，可以抑制振荡和不稳定趋向利温伯格对λ和D的选择提出若干建议。其中之一是取

$$\lambda^{(k)}=\frac{e^{T(k)}X^{(k)}X^{T(k)}e^{(k)}}{Q^{(k)}} \tag{3-3-49}$$

这时λ 随着Q逐步接近最小而改变，由于Q→min时，X^Te的每个分量→0，而λ 的分子较分母Q更快地趋于零，故 $\lambda^{(k)}$随着$Q^{(k)}$趋于零。

麦夸特(*Marguardt*)的方法是在式(3-3-48)中取

$$\lambda^{(k)}=\lambda_0/c^k \tag{3-3-50}$$

c为大于1的常数。

$$D=\begin{pmatrix} \sum\left(\frac{\partial f_t}{\partial\beta_1}\right)^2 & & O \\ & \ddots & \\ O & & \sum\left(\frac{\partial f_t}{\partial\beta_q}\right)^2 \end{pmatrix} \tag{3-3-51}$$

$=X^TX$的主对角线元素

这是为了避免参数比例尺不同的影响，而且可以保持步长系数中量纲的一致性。麦夸特方法在迭代开始阶段接近于最速下降法，而在接近迭代终点时近似于高斯牛顿法。具有式(3-3-48)形式的算法统称阻尼高斯一牛顿法，它是最速下降法和高斯一牛顿法的折衷。

在实用程序中，规定了第一次计算的λ_0值，以后如果计算结果表明新估计参数对应的残差平方和减小了，则将原来的λ除以某常数c作为下一次计算的λ值；相反，如果计算结果发现残差平方和无改善甚至更大，则将λ加大若干倍重新计算。因此在整个迭代过程中，后一步所用的λ比前一步来说可能是有增有减的。

参数估计的置信区间可以利用Cov$[\beta_{LS}]$的主对角线元素，即估计值的方差来计算。如果采用阻尼高斯一牛顿法，（最后一

次迭代的步长为$\hat{\beta}_{LS}-\beta_0$)，则近似取为

$$\begin{aligned}\mathrm{Cov}[\hat{\beta}_{LS}]&=E[(\hat{\beta}_{LS}-\beta_0)(\hat{\beta}_{LS}-\beta_0)^T]\\&=E[(X^T(\beta_0)X(\beta_0)+\lambda D)^{-1}X^T(\beta_0)ee^TX(\beta_0)\\&\quad(X^T(\beta_0)X(\beta_0)+\lambda D)^{-1}]\\&=\sigma_\epsilon^2AX^{-1}(\beta_0)X(\beta_0)A\end{aligned}$$

其中$A=[X^T(\beta_0)X(\beta_0)+\lambda D]^{-1}$，且设$E[ee^T]=\sigma_\epsilon^2\mathrm{I}$。通常取$\pm1.96\times$(估计的标准差)作为估计的95%置信区间。

附录中给出的MARQT计算机源程序可供使用参考。

3.4 模型参数的极大似然估计(ML估计)

1. 极大似然估计 费歇(*Fisher*)提出的极大似然估计是一种具有普遍性的有效的估计方法，只要能够写出观察值的联合概率分布，就可用它来解决估计问题。

设$\{x_t\}$为一随机序列，用$p(x_1, \cdots, x_N)$表示其有穷维概率密度。假定该概率密度能被一组参数$\alpha=(\alpha_1, \alpha_2, \cdots\alpha_q)^T$唯一确定，我们就用$p(x_1, \cdots, x_N|\alpha)$记在$\alpha$给定条件下$x_1, x_2, \cdots x_N$的概率密度，称之为条件概率密度，这个量描述了以$\alpha$为参数时，样本值$x_1, \cdots, x_N$出现的可能性，很自然地希望选择参数$\alpha$以使样本值$x_1, \cdots, x_N$出现的可能性最大。因此，我们相对于$\alpha$极大化$p(x_1, \cdots, x_N|\alpha)$，使之达到极大的记作$p(x_1, \cdots, \alpha_N|)\alpha$。这就是说，以$\alpha$为参数时，样本值$x_1, \cdots, x_N$出现的可能性最大。由于$x_1, \cdots, x_N$是已得到的样本，我们可以把它看做是固定的量值，而把α看作是待定参数。为突出这一对应关系，令

$$L(\alpha|x_1, \cdots, x_N)=p(x_1, \cdots, x_N|\alpha) \qquad (3\text{-}4\text{-}1)$$

我们称L为$x_1, \cdots, x_N$的似然函数。上述极大化概率密度，就变成求似然函数$L(\alpha|x_1, \cdots, x_N)$极大值的问题了。使$L(\alpha|x_1, \cdots, x_N)$达到极大的$\alpha$称为极大似然估计。

若$\{x_t\}$是正态过程，那么$p(x_1, \cdots, x_N|\alpha)$便是指数型函数，由于对数函数是单调增加的，因此在求解极值问题时，采用以下等价形式在计算上更为方便：

$$l(\alpha|x_1, \cdots, x_N)=\log L(\alpha|x_1, \cdots, x_N)$$

l称为$x_1, \cdots, x_N$的对数似然函数。于是求极大似然估计问题又化成求解如下方程：

$$\left.\frac{\partial l}{\partial \alpha}\right|_{\alpha=\hat{\alpha}}=0 \qquad (-4-2)$$

该方程称为似然方程，它的解$\hat{\alpha}$称为参数α的极大似然估计。

2. AR、MA和ARMA模型参数的极大似然估计 设$\{x_t\}$是零均值、正态$ARMA(n, m)$序列，其概率密度被模型参数$a_1, \cdots, a_n, b_1, \cdots, b_m$及$\sigma_\varepsilon^2$唯一确定。引入参数向量记号

$$a=[a_1, \cdots, a_n]^T, \qquad b=[b_1, \cdots, b_m]^T$$

$$\beta^T=[a^T, b^T], \qquad \alpha^T=[\beta^T, \sigma_\varepsilon^2]$$

记$X_N=(x_1, x_2, \cdots, x_N)^T$，则$(x_1, \cdots x_N)$的概率密度为

$$p(X_N|\alpha)=(2\pi\sigma_\varepsilon^2)^{-\frac{N}{2}}|M_N|^{\frac{1}{2}}\exp\left\{-\frac{X_N^T M_N X_N}{2\sigma_\varepsilon^2}\right\} \qquad (3-4-3)$$

其中$M_N\equiv\sigma_\varepsilon^2 R_N^{-1}$而$R_N=E[X_N X_N^T]$是样本协方差阵。对数似然函数可以写成

$$l(\alpha|x_1, \cdots, x_N)=-\frac{N}{2}\log 2\pi+\frac{1}{2}\log|M_N|-\frac{N}{2}\log\sigma_\varepsilon^2$$

$$-\frac{X_N^T M_N X_N}{2\sigma_\varepsilon^2} \qquad (3-4-4)$$

由于公式复杂，无法得出求解极限的解析表示，用数值法求解也过于繁复。事实上，很少采用严格的极大似然估计，而是用近似的估计值来代替。

由式(3-4-4)可以看出右边第一项只与N有关，而和参数α无关。另外，可以证明（读者可参阅([4]书后附录§3)，右边第二项中的$|M_N|$具有与N无关的上界，而后两项是和N同阶的，因此，只要N充分大时，式(3-4-4)是否达到极大值取决于右边后两项：$-\frac{N}{2}\log\sigma_\varepsilon^2-\frac{X_N^T M_N X_N}{2\sigma_\varepsilon^2}$，换句话说，$l(\alpha|x_1,\cdots,x_N)$取极大值的点，当N充分大时与

$$\frac{N}{2}\log\sigma_\varepsilon^2+\frac{X_N^T M_N X_N}{2\sigma_\varepsilon^2} \tag{3-4-5}$$

取极小值的点几乎是一样的。对式(3-4-5)求极小而得到的估计值$\hat{\alpha}$，就是在给定样本$x_1,\cdots x_N$时模型ARMA(n, m)参数的近似极大似然估计。

在给出近似极大似然估计的表达式之前，需要对$X_N^T M_N X_N$作一点说明：$X_N^T M_N X_N$仅与参数β有关而和σ_ε^2无关，通常将它记作$Q(\beta)$，并称其为平方和函数，因为

$$Q(\beta)=X_N^T M_N X_N=\sum_{t=1}^{N}\hat{\varepsilon}_t^2 \tag{3-4-6}$$

其中$\hat{\varepsilon}_t$是在给定参数值和有穷个样本值$(x_1, x_2\cdots, x_N)$下ε_t的最小方差估计，记作

$$\hat{\varepsilon}_t=E_\beta(\varepsilon_t|X_N)=E_\beta(\varepsilon_t|x_1,\ \cdots,\ x_N) \tag{3-4-7}$$

关于线性最小方差估计的基本概念是这样的，设观察值X和被估计量Y的概率密度未知，为简化问题，限定所求的估计量$\hat{Y}$必须是观察值X的线性函数：

$$\hat{Y}=a+BX \tag{3-4-8}$$

其中a是非随机向量，B是常值矩阵，求线性最小方差估计就是要选择a和B使得

$$E(Y-\widehat{Y})(Y-\widehat{Y})^{\mathrm{T}}=E(Y-a-BX)(Y-a-BX)^{\mathrm{T}} \tag{3-4-9}$$

达到极小。为求出这种最优估计$\widehat{a}$和$\widehat{B}$，令

$$b=a+BEX-EY \tag{3-4-10}$$

于是式(3-4-9)可写成

$$\begin{aligned}E(Y-\widehat{Y})(Y-\widehat{Y})^{\mathrm{T}}&=E[(Y-EY)-b-B(X-EX)][(Y-EY)\\&\quad-b-B(X-EX)]^{\mathrm{T}}\\&=\mathrm{Var}Y+bb^{\mathrm{T}}+B(\mathrm{Var}X)B^{\mathrm{T}}-\mathrm{Cov}(Y,X)B^{\mathrm{T}}-B\mathrm{Cov}(X,Y)\\&=bb^{\mathrm{T}}+[B-\mathrm{Cov}(Y,X)(\mathrm{Var}X)^{-1}](\mathrm{Var}X)[B-\mathrm{Cov}(Y,X)\\&\quad(\mathrm{Var}X)^{-1})]^{\mathrm{T}}+[\mathrm{Var}Y-\mathrm{Cov}(Y,X)(\mathrm{Var}X)^{-1}\mathrm{Cov}(Y,X)^{\mathrm{T}}]\end{aligned} \tag{3-4-11}$$

该式右边前两项都是非负定阵，第三项与B无关，显然要使式(3-4-11)达到极小，就要选择b和B使该式中前两项等于零，即

$$\widehat{b}=0, \quad \widehat{B}=\mathrm{Cov}(Y,X)(\mathrm{Var}X)^{-1} \tag{3-4-12}$$

将式(3-4-12)代入式(3-4-10)得

$$\widehat{a}=EY-\mathrm{Cov}(Y,X)(\mathrm{Var}X)^{-1}EX \tag{3-4-13}$$

从而得到线性最小方差估计为

$$\widehat{Y}=EY+\mathrm{Cov}(Y,X)(\mathrm{Var}X)^{-1}(X-EX) \tag{3-4-14}$$

现在假定$\{x_t\}$，$\{\varepsilon_t\}$都是零均值，在任何t有$Ex_t=0$，$E\varepsilon=0$，则利用式(3-4-14)可得ε_t的最小方差估计

$$\begin{aligned}\widehat{\varepsilon}_t&=E_t(\varepsilon_t|X_N)=[E(\varepsilon_tX_N)]^{\mathrm{T}}[E(X_NX_N^{\mathrm{T}})]^{-1}X_N\\&=[E(\varepsilon_tX_N)]^{\mathrm{T}}R_N^{-1}X_N\end{aligned} \tag{3-4-15}$$

因此

$$\sum_{t=-\infty}^{N}\widehat{\varepsilon}_t^2=\sum_{t=-\infty}^{N}X_N^{\mathrm{T}}R_N^{-1}E(\varepsilon_tX_N)E(\varepsilon_tX_N^{\mathrm{T}})R_N^{-1}X_N$$

$$=X_N^{\mathrm{T}}R_N^{-1}\sum_{t=-\infty}^{N}E(\varepsilon_tX_N)E(\varepsilon_tX_N^{\mathrm{T}})R_N^{-1}X_N \tag{3-4-16}$$

利用模型的传递形式，可以写出

$$E(\varepsilon_t x_k)=\sum_{j=0}^{\infty}G_j E(\varepsilon_t \varepsilon_{k-j})=\begin{cases}\sigma_\varepsilon^2 G_{k-t} & t\leqslant k\\ O & t>k\end{cases}$$

再利用式(2-3-74)可以写出

$$R(i-j)=\sigma_\varepsilon^2\sum_{k=0}^{\infty}G_k G_{k+j-i}=\sigma_\varepsilon^2\sum_{t=-\infty}^{N}G_{i-t}G_{j-t}$$

这里形式地记$G_i=0$　($i<0$)。从而有

$$\sum_{t=-\infty}^{N}E(\varepsilon_t X_N)E(\varepsilon_t X_N)^{\mathrm{T}}=\sigma_\varepsilon^4\begin{pmatrix}G_{1-t}\\ G_{2-t}\\ \vdots\\ G_{N-t}\end{pmatrix}(G_{1-t}\ G_{2-t},\cdots,G_{N-t})$$

$$=\sigma_\varepsilon^2 R_N$$

代入式(3-4-16)得

$$\sum_{t=-\infty}^{N}\varepsilon_t^2=X_N^{\mathrm{T}}R_N^{-1}R_N R_N^{-1}X_N\sigma_\varepsilon^2=\sigma_\varepsilon^2 X_N^{\mathrm{T}}R_N^{-1}X_N$$

$$=X_N^{\mathrm{T}}M_N X_N$$

这就是式(3-4-6)。由于$\sigma_\varepsilon^2 R_N$仅由格林函数决定，而后者只取决于模型参数$\boldsymbol{\beta}$；因此$M_N(=\sigma_\varepsilon^2 R_N^{-1})$只是$\boldsymbol{\beta}$的函数。此外，$X_N$是已得到的样本值，所以$X_N^{\mathrm{T}}M_N X_N$也仅是$\boldsymbol{\beta}$的函数，记作$Q(\boldsymbol{\beta})$并称之为平方和函数。

在式(3-4-5)中对参数$\boldsymbol{\alpha}=(\boldsymbol{\beta}^{\mathrm{T}},\ \sigma_\varepsilon^2)$求偏导数，得方程组

$$\frac{\partial}{\partial\sigma_\varepsilon^2}\left(\frac{N}{2}\log\sigma_\varepsilon^2+\frac{Q(\boldsymbol{\beta})}{2\sigma_\varepsilon^2}\right)\Bigg|_{\boldsymbol{\alpha}=\widehat{\boldsymbol{\alpha}}}=\frac{N}{2\hat{\sigma}_\varepsilon^2}-\frac{Q(\widehat{\boldsymbol{\beta}})}{2\hat{\sigma}_\varepsilon^4}=0$$

$$\frac{\partial}{\partial\boldsymbol{\beta}}\left(\frac{N}{2}\log\sigma_\varepsilon^2+\frac{Q(\boldsymbol{\beta})}{2\sigma_\varepsilon^2}\right)\Bigg|_{\boldsymbol{\alpha}=\widehat{\boldsymbol{\alpha}}}=\frac{1}{2\sigma_\varepsilon^2}\frac{\partial}{\partial\boldsymbol{\beta}}Q(\boldsymbol{\beta})\Bigg|_{\boldsymbol{\beta}=\widehat{\boldsymbol{\beta}}}=0$$

由此可得

$$\widehat{\sigma}_\varepsilon^2=\frac{1}{N}Q(\widehat{\boldsymbol{\beta}})$$

而$\widehat{\boldsymbol{\beta}}$是方程

$$\left.\frac{\partial Q(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right|_{\boldsymbol{\beta}=\widehat{\boldsymbol{\beta}}} = 0 \tag{3-4-8}$$

的解。由此我们得到了使$Q(\widehat{\boldsymbol{\beta}})$达极小的参数估计值$\widehat{\boldsymbol{\beta}}$和$\sigma_\varepsilon^2$，当样本充分大时，该估计近似地等于参数的极大似然估计。

适合于方程式(3-4-18)的参数$\widehat{\boldsymbol{\beta}}$使平方和函数$Q(\boldsymbol{\beta})$达到极小，我们称之为最小平方和估计。根据前述，当样本充分大时，对正态$ARMA$序列而言，参数的最小平方和估计是极大似然估计的极好近似。

下面我们针对时间序列的各种情况，进一步讨论平方和函数$Q(\boldsymbol{\beta})$的计算。

对于AR序列，我们可以给出平方和函数关于模型参数的解析表示式，对于MA或$ARMA$序列，则不能得出Q关于参数$\boldsymbol{\beta}$的解析表达式，因而只能通过迭代计算实现平方和函数的极小化。

(1)AR(n)序列最小平方和估计及$Q(\boldsymbol{a})$的计算。对AR(n)序列，参数向量$\boldsymbol{\beta}=\boldsymbol{a}$，经过繁杂的计算可以得到，当$N>2n$时，平方和函数的解析表示是

$$Q(\boldsymbol{a}) = \sum_{i;j=0}^{n} a_i a_j \tilde{R}_{ij}, \qquad N \geqslant 2n \tag{3-4-19}$$

其中 $a_0=1$， $\tilde{R}_{ij}=\sum_{t=1+j}^{N-i} x_{t+i-j} x_t = \tilde{R}_{ji}$ $\quad i,j=0,1,\cdots$

在实际问题中$N>2n$通常都可以做到。

记

$$\tilde{R}=\begin{pmatrix} \tilde{R}_{11} & \tilde{R}_{12} & \cdots & \tilde{R}_{1n} \\ \tilde{R}_{21} & \tilde{R}_{22} & \cdots & \tilde{R}_{2n} \\ \cdots & & & \\ \tilde{R}_{n1} & \tilde{R}_{n2} & \cdots & \tilde{R}_{nn} \end{pmatrix} \quad \tilde{R}_0=(\tilde{R}_{01}, \tilde{R}_{02}, \cdots \tilde{R}_{0n})^T$$

(3-4-19)式又可以表为

$$Q(a)=a^{T}\tilde{R}a-2a^{T}\tilde{R}_{0}+\tilde{R}_{00} \tag{3-4-20}$$

这是参数a的二次型函数，使$Q(a)$达到极小的解，是参数a的最小平方和估计：

$$\widehat{a}_{Q}=\widetilde{R}^{-1}\widetilde{R}_{0}$$

σ_{ε}^{2}的最小平方和估计由式(3-4-17)得

$$\sigma_{\varepsilon}^{2}=\frac{1}{N}Q(\widehat{a}_{Q})=\frac{1}{N}(\widetilde{R}_{00}-\widetilde{R}_{0}^{T}\widetilde{R}^{-1}\widetilde{R}_{0})$$

我们来考察一下最小平方和估计与最小二乘估计之间的关系。当N足够大时有

$$N\tilde{R}_{ij}=\sum_{t=1+i}^{N-i}x_{t+i-j}x_{t}\approx\sum_{t=n+1}^{N}x_{t+n-i}x_{t+n-j}=N\hat{R}'(i-j)$$

上式最右端和第3.3中式(3-3-18)的定义是一致的。这表明当样本个数N足够大时，$AR(n)$模型参数的最小平方和估计与最小二乘估计是十分接近的。而在节3.3中已经讲过，对充分大的N，矩估计与最小二乘估计很相近。因此，对足够大的样本，前述三种估计都是很接近的。

对于AR序列而言，平方和函数除了可用式(3-4-19)进行计算，还可以直接由模型参数及样本计算，读者可参看(4)中附录的有关部分。

(2)ARMA序列最小平方和函数的计算

在MA和$ARMA$序列中，参数β与观察值是非线性关系，因此，计算$Q(\beta)$比$AR(n)$序列复杂得多，而且只能得到$Q(\beta)$的近似值。

$ARMA$序列平方和函数的计算公式仍然是式(3-4-6)，其中的ε_{t}应是由已有的样本值X_{N}得到的线性最小方差估计。当样本个数N充分大时，为计算简便可近似地取

$$\hat{\varepsilon}_t = x_t - \hat{a}_1 x_{t-1} - \cdots - \hat{a}_n x_{t-n} + \hat{b}_1 \hat{\varepsilon}_{t-1} + \cdots + \hat{b}_m \hat{\varepsilon}_{t-m} \qquad (3\text{-}4\text{-}21)$$

并约定

$$\hat{\varepsilon}_t = 0 \qquad t \leqslant n \qquad (3\text{-}4\text{-}22)$$

当t趋无穷时，由式(3-4-21)递推得到的$\hat{\varepsilon}_t$与ε_t的最小方差估计十分接近（参见[4]）。由于式(3-4-22)中$\hat{\varepsilon}_t$的初值取为零，因此用有穷和代替无穷和，于是有

$$Q(\beta) = \sum_{t=n+1}^{N} \hat{\varepsilon}_t \qquad (3\text{-}4\text{-}23)$$

式(3-4-23)给出的近似平方和函数$Q(\beta)$，是参数β的非线性函数，必须通过非线性最小二乘迭代求近似最优解。

3.5 模型阶数的确定

前面讨论模型的参数估计都是假设自回归和滑动平均的阶数事先已知，然而事实上这些阶数几乎总是不知道的。它们应当是另外的待估参数，其值需根据观察数据来做适当的判定。

1. **残差方差图** 在AR模型的情况下，很自然地会想到按照统计学中多元回归分析的方法，通过假设检验来确定阶数。也就是以AR(n)模型作为零假设，而以AR($n+1$)模型作为对立假设，然后增大n值，直到检验结果不显著为止。最早对这类检验进行研究的是惠多(*Whittle*)。

假定模型是有限阶的自回归模型，如果选择的n小于真正的阶数，则$\hat{\sigma}_\varepsilon^2$将比真正模型的残差方差$\sigma_\varepsilon^2$大。这是因为我们把模型中本来应有的一些高阶项略去了，而这些项对于减小残差方差是有明显贡献的。另一方面，如果n已经达到真值，那么再进一步增加阶数，并不会使$\hat{\sigma}_\varepsilon^2$有显著减小，甚至还略有增加。

这样用一系列阶数递增的模型进行拟合，每次都求出$\widehat{\sigma_\epsilon^2}$，然后画$n$和$\hat{\sigma}_\epsilon^2$的图形——称为残差方差图。开始时$\hat{\sigma}_\epsilon^2$会下降，当达到$n$的真值后渐趋平缓。詹金斯(*Jenkins*)和瓦茨(*Watts*)曾使用这种方法，并且也用来确定*MA*及*ARMA*模型的阶数。残差方差的估计式为：

$$\text{AR模型}\ \hat{\sigma}_\epsilon^2(n)=\frac{1}{N-2n-1}Q(\hat{\mu},\hat{a}_1,\cdots,\hat{a}_n)$$

$$MA\text{模型}\ \hat{\sigma}_\epsilon^2(m)=\frac{1}{N-m-1}Q(\hat{\mu},\hat{b},\cdots,\hat{b}_m)$$

$$\text{ARMA模型}\ \hat{\sigma}_\epsilon^2(n,m)=\frac{1}{N-2n-m-1}Q(\hat{\mu},\ \hat{a}_1,\ \cdots,\ \hat{a}_n,\hat{b}_1,\cdots,\hat{b}_m)$$

这里当$n\neq 0$时实际观察数取$N-n$个。

图3-2是对某化学蒸馏塔每批输出量的记录数(共70个数)，用不同阶数的ARMA模型进行拟合，图3-3是根据MARQT程序计算结果画出的残差方差图(计算中不作μ的LS估计)由图可见，

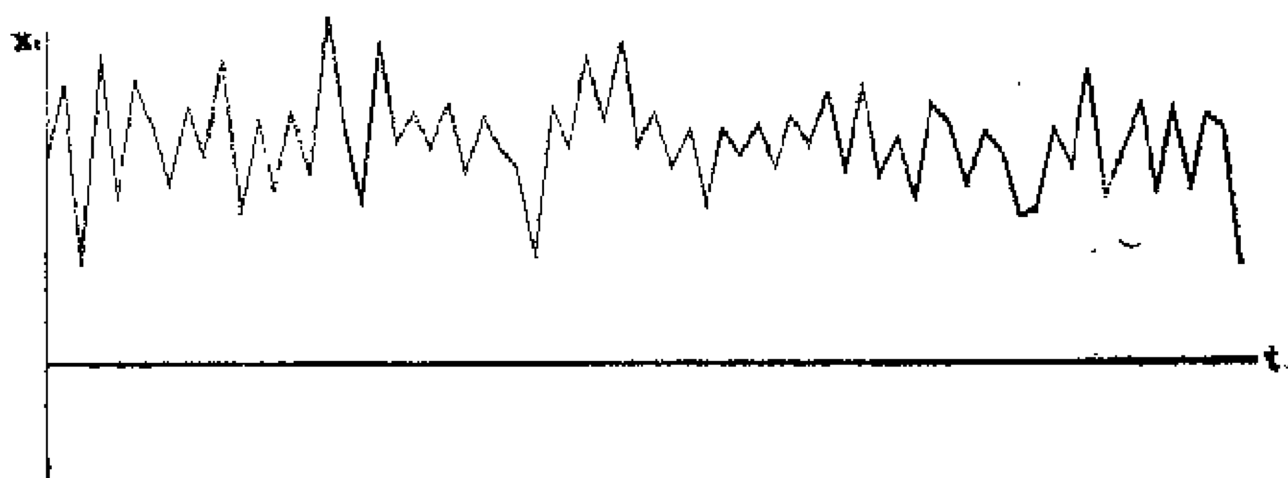


图3-2 蒸馏塔批输出量数据图

在$n+m$小于8的情况下AR(2)模型的拟合残差方差最小。从总的趋势看，残差方差随着阶数的增加是减小的。

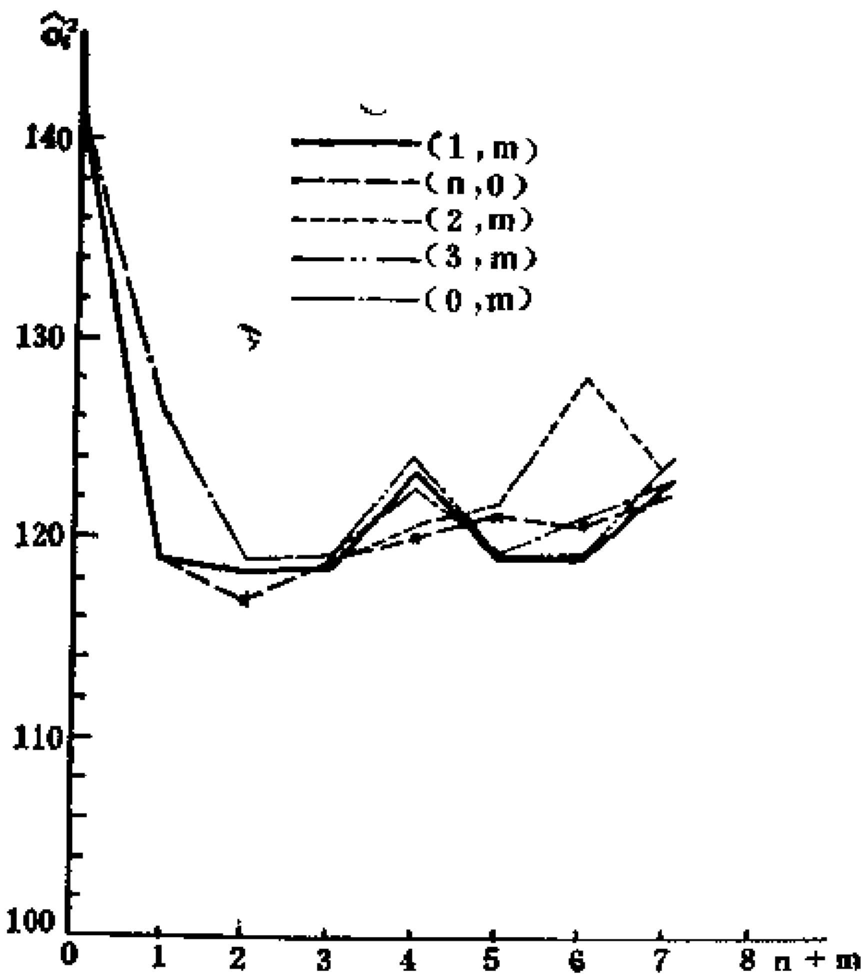


图3-3 蒸馏塔数据模型的残差方差图

2. 偏相关函数和模型 波克斯（Box）和詹金斯(Jenkins)建议在AR模型的情况下，用偏相关函数来提供有关阶数的信息，因为由AR(k)的最后一个系数可得x_{t-k}和x_t之间的偏相关系数（x_{t-k+1}，$\cdots x_{t-1}$固定时），如果真正模型是AR(n)，则当$k>n$时偏相关为零，因此，做出偏相关系数对于k的关系曲线可以用来估计真实模型的阶数。

下面首先对偏相关函数的提出及其含义作简略的说明。

在第二章介绍AR模型的自相关函数时可以看出，尽管它只涉及n步的相关性，但自相关函数要到$r\to\infty$时才衰减到零。或者说它的自相关函数是拖尾的。有人可能会感到奇怪，既然AR(1)模型只涉及"一步相关性"为什么$\rho(r)$在$r=1$之后不截尾，而当$r\to\infty$才逐渐趋于零？事实上，这是因为x_t和x_{t-1}有关，而x_{t-1}又和x_{t-2}有关。所以x_t和x_{t-2}有关，进一步，x_{t-2}又和x_{t-3}有关，如此类推，x_t和x_{t-3}，$x_{t-4}\cdots$也都有关系。这种情况说明了用简单的相关函数来说明两个变量之间内在的相关性，有时会导致错误的印象。因为假如两个变量之间有正的自相关，其原因可能有两种：一是一个变量的增加导致另一变量的增加；另一可能是某一个（或多个）变量的变化使被考察的两个变量都增加了。因此需要有另一种表达两个变量之间相关性的方法，它应当能够排除其它（一个或多个）变量的影响。这就是偏相关函数，它对于模型类的估计也将有指导意义。

可以设想，如果在排除了"局外"变量的影响之后，两个变量本身实际上并不存在相关性，就说明看起来似乎密切的相互关系是由于其它变量作用的缘故。反之，假若两个变量的相关值小于偏相关值，那么必定是其它变量"冲淡"了这两个变量间的相互关系。

根据以上所述，如果有三个随机变量u,v,w，其联合概率密度为$p(u,v,w)$，在w给定条件下u和v的条件概率密度记作$p(u,$

$v|w)=p(u,v,w)/p(w)$，则w已给时，u和v的偏相关应定义为

$$\rho_{u,v,w}=\frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(u-\bar{u})(v-\bar{v})p(u,v|w)dudv}{(\mathrm{Var}(u)\mathrm{Var}(v))^{1/2}}$$

或

$$\rho_{u,v,w}=\frac{\underset{w}{E}((u-\bar{u})(v-\bar{v}))}{(\mathrm{Var}(u)\mathrm{Var}(v))^{1/2}}$$

其中$\bar{u}=E(u)$，$\bar{v}=E(v)$，而$\underset{w}{E}$表示关于条件密度函数$p(u,v|w)$的条件期望。

在时间序列模型中，为了方便，定义k步迟后的偏相关为k－1个中间值x_{t-1}，x_{t-2}，$\cdots x_{t-(k-1)}$一定时，x_t和x_{t-k}之间的偏相关。其记号和定义式为

$$\rho_{x_t x_{t-k}\cdot x_{t-1}x_{t-2}\cdots x_{t-k+1}}=\frac{E(x_t x_{t-k})}{\sqrt{E(x_t^2)E(x_{t-k}^2)}}$$

$$=\frac{E(x_t x_{t-k})}{\mathrm{Var}(x_t)}$$

这里的E是关于条件密度函数$p(x_t,x_{t-k}|x_{t-1},\cdots,\ x_{t-k+1})$的条件期望。

下面来看AR(k)过程的偏相关函数。设x_t的均值为零，它所适合的模型为

$$x_t=\varphi_{k1}x_{t-1}+\varphi_{k2}x_{t-2}++\cdots+\varphi_{kk}x_{t-k}+\varepsilon_t$$

在上式两边乘以x_{t-k}（$k>0$）并取条件期望，即

$$E(x_t x_{t-k})=\varphi_{k1}E(x_{t-1}x_{t-k})+\varphi_{k2}E(x_{t-2}x_{t-k})+\cdots$$
$$+\varphi_{kk}E(x_{t-k}^2)+E(\varepsilon_t x_{t-k})$$

因$x_{t-1},\cdots,\ x_{t-k+1}$是既定条件，可提到条件期望运算符号之外，即

$$E(x_t x_{t-k})=\varphi_{k1}x_{t-1}(Ex_{t-k})+\varphi_{k2}x_{t-2}E(x_{t-k})+\cdots\cdots$$

$$+\varphi_{kk}x_{t-k+1}E(x_{t-k})+\varphi_{kk}E(x_{t-k}^2)+E(\varepsilon_t x_{t-k})$$

由于x_t是均值为零的平稳过程，$E(x_{t-1})=E(x_{t-2})=\cdots=E(x_{t-k+1})=0$，$E(x_{t-k}^2)=Var(x_t)$，于是

$$E(x_t x_{t-k})=\varphi_{kk}Var(x_t)+E(\varepsilon_t x_{t-k})$$

当k>0时ε_t和x_{t-k}无关，上式第二项为零。故有

$$\varphi_{kk}=\frac{E(x_t x_{t-k})}{Var(x_t)}$$

因此当迟后为k时，AR(k)过程的偏相关函数值φ_{kk}，就是该AR(k)模型的最后一个自回归系数的负数（即$-a_k$）。

下面来看偏相关函数和自相关函数的关系。以便从$\rho(r)$求偏相关 $\{\varphi_{11},\varphi_{22},\cdots,\varphi_{kk},\cdots\}$

对于k阶自回归过程，其耶尔一瓦克尔方程为

$$\rho(r)-\varphi_{k1}\rho(r-1)-\varphi_{k2}\rho(r-2)-\cdots-\varphi_{kk}\rho(r-k)=0$$

$$r=1,2,3,\cdots \qquad (3\text{-}5\text{-}1)$$

当$k=0$时得

$$\varphi_{00}=\rho(0)=1 \qquad (3\text{-}5\text{-}2)$$

当$k=1$时，得

$$\rho(r)-\varphi_{11}\rho(r-1)=0 \qquad r=1,2,3\cdots,$$

$$\varphi_{11}=\rho(r)/\rho(r-1) \qquad r=1,2,3\cdots,$$

取$r=1$，有

$$\varphi_{11}=\rho(1)/\rho(0)=\rho(1) \qquad (3\text{-}5\text{-}3)$$

当$k=2$时，得

$$\rho(r)-\varphi_{21}\rho(r-1)-\varphi_{22}\rho(r-2)=0 \quad (r=1,2,3,\cdots)$$

取$r=1$和$r=2$，得两个方程式

$$\rho(1)-\varphi_{21}\rho(0)-\varphi_{22}\rho(1)=0$$

$$\rho(2)-\varphi_{21}\rho(1)-\varphi_{22}\rho(0)=0$$

解得

$$\varphi_{22}=\frac{\rho(2)\rho(0)-\rho^2(1)}{\rho^2(0)-\rho^2(1)}=\frac{\rho(2)-\rho^2(1)}{1-\rho^2(1)} \tag{3-5-4}$$

对于一般情况的耶尔-瓦克尔方程如(3-5-1)，取$r=1,2,\cdots,k$得k个方程

$$\begin{aligned}
&\rho(1)-\varphi_{k1}\rho(0)-\varphi_{k2}\rho(1)-\cdots-\varphi_{kk}\rho(k-1)=0\\
&\rho(2)-\varphi_{k1}\rho(1)-\varphi_{k2}\rho(0)-\cdots-\varphi_{kk}\rho(k-2)=0\\
&\cdots\\
&\rho(k)-\varphi_{k1}\rho(k-1)-\varphi_{k2}\rho(k-2)-\cdots-\varphi_{kk}\rho(0)=0
\end{aligned} \tag{3-5-5}$$

或写成矩阵形式

$$\begin{pmatrix} 1 & \rho(1) & \cdots & \rho(k-1)\\ \rho(1) & 1 & \cdots & \rho(k-2)\\ \cdots & & & \\ \rho(k-1) & \rho(k-2) & \cdots & 1\end{pmatrix}\begin{pmatrix}\varphi_{k1}\\ \varphi_{k2}\\ \vdots\\ \varphi_{kk}\end{pmatrix}=\begin{pmatrix}\rho(1)\\ \rho(2)\\ \vdots\\ \rho(k)\end{pmatrix} \tag{3-5-6}$$

用求解线性方程组的克莱姆（*Cramer*）法则可解得

$$\varphi_{kk}=\begin{vmatrix} 1 & \rho(1) & \cdots & \rho(1)\\ \rho(1) & 1 & \cdots & \rho(2)\\ \cdots & & & \\ \rho(k-1) & \rho(k-2) & \cdots & \rho(k)\end{vmatrix}\Bigg/\begin{vmatrix} 1 & \rho(1) & \cdots & \rho(k-1)\\ \rho(1) & 1 & \cdots & \rho(k-2)\\ \cdots & & & \\ \rho(k-1) & \rho(k-2) & \cdots & 1\end{vmatrix} \tag{3-5-7}$$

例如考察$AR(1)$过程$x_t+ax_{t-1}=\varepsilon_t$，在$x_{t-1}$已给条件下，$x_t$和$x_{t-2}$的偏相关是

$$\rho_{x_t x_{t-2}\cdot x_{t-1}}=\frac{\rho(2)-\rho(1)\rho(1)}{1-\rho(1)\rho(1)}=\frac{a^2-a.a}{1-a.a}=0$$

这里利用了AR过程自相关的公式$\rho(r)=a^r$。可见尽管$\rho(2)=a^2$，但x_t和x_{t-2}之间的关系是完全通过x_{t-1}而产生的。如果排除了x_{t-1}的影响，则x_t和x_{t-2}是不相关的。当$r>1$时偏相关也是零。这就说明了该过程的一步相关性。

对于AR(n)过程来说，偏相关函数 $\{\varphi_{kk}\}$ 在$k>n$以后全为零。这种截尾性是AR过程特有的属性。由于MA或ARMA过程都相当于阶数无限的 AR 过程，显然其偏相关函数将是拖尾的。下表列出不同类型过程的自相关和偏相关的尾部性质以作比较。这些性质将有助于估计模型的类别和阶数。

	AR (n)	MA (m)	ARMA (n,m)
自相关函数	拖尾	截尾	拖尾
偏相关函数	截尾	拖尾	拖尾

图3-4为以下过程对应的偏相关函数的图形

AR:

(1) $x_t+0.8x_{t-1}=\varepsilon_t$

(2) $x_t-0.8x_{t-1}=\varepsilon_t$

(3) $x_t+0.5x_{t-1}-0.3x_{t-2}=\varepsilon_t$

(4) $x_t-0.5x_{t-1}-0.3x_{t-2}=\varepsilon_t$

MA:

(5) $x_t=\varepsilon_t-0.8\varepsilon_{t-1}$

(6) $x_t=\varepsilon_t+0.8\varepsilon_{t-1}$

(8) $x_t=\varepsilon_t+0.5\varepsilon_{t-1}-0.3\varepsilon_{t-2}$

(8) $x_t=\varepsilon_t-0.5\varepsilon_{t-1}-0.3\varepsilon_{t-2}$

$ARMA$:

(9) $x_t+0.7x_{t-1}=\varepsilon_t+0.4\varepsilon_{t-1}$

(10) $x_t+0.4x_{t-1}=\varepsilon_t+0.7\varepsilon_{t-1}$

(11) $x_t-0.5x_{t-1}=\varepsilon_t+0.5\varepsilon_{t-1}$

(12) $x_t+0.5x_{t-1}=\varepsilon_t-0.5\varepsilon_{t-1}$

实际计算中往往使用以下递推公式求偏相关函数的值（程序可见附录二PARCR）：

$\varphi_{11}=\rho(1)$

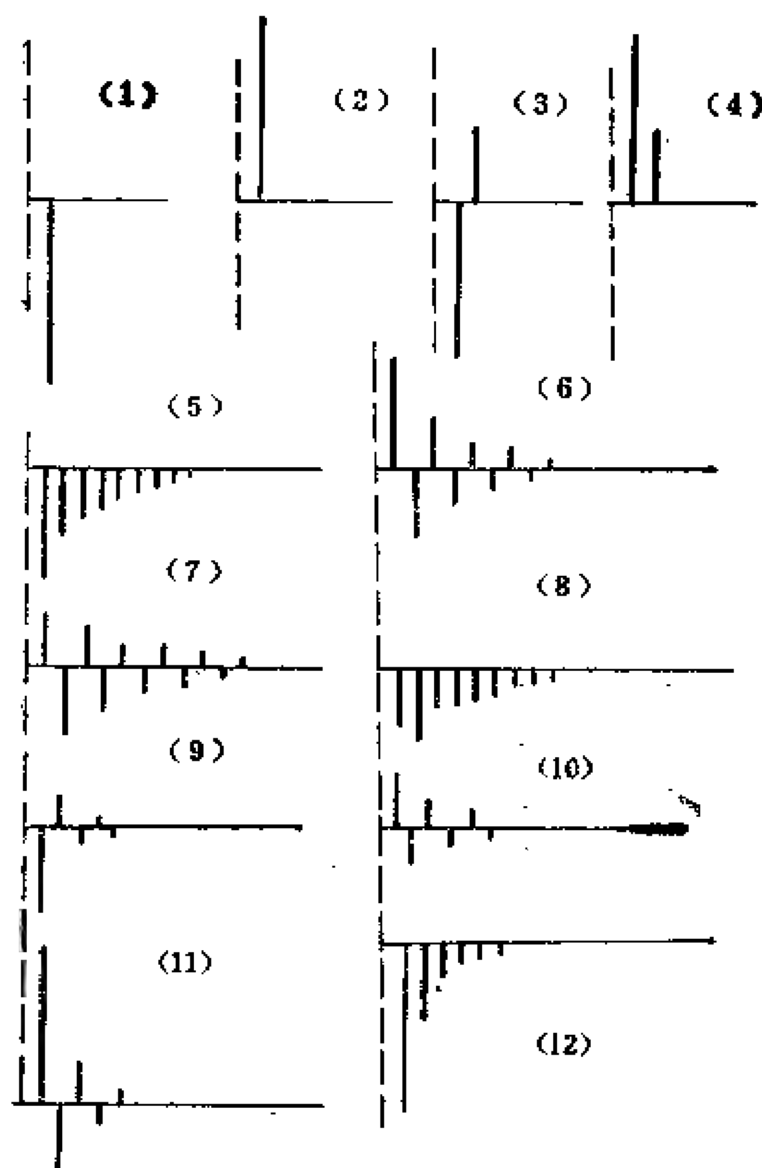


图3-4　偏相关函数图形

$$\varphi_{k+1,\ j}=\varphi_{k\ j}-\varphi_{k+1,\ k+1}\varphi_{k,\ k-(j-1)} \tag{3-5-8}$$

$$\varphi_{k+1,\ k+1}=\left[\rho(k+1)-\sum_{j=1}^{k}\rho(k+1-j)\varphi_{k\ j}\right]/\left[1-\sum_{j=1}^{k}\rho(j)\varphi_{k\ j}\right]$$

$$j=1,2,\cdots k.$$

该式推导如下，令，

$$\boldsymbol{\rho}_k=\begin{pmatrix}1 & \rho(1) & \cdots & \rho(k-1)\\ \rho(1) & 1 & \cdots & \rho(k-2)\\ \cdots & & & \\ \rho(k-1) & \rho(k-2) & \cdots & 1\end{pmatrix},\qquad T=\begin{pmatrix}0 & & & 1\\ & & \cdot^{\cdot^{\cdot}} & \\ & & 1 & \\ 1 & & & 0\end{pmatrix},$$

$$\boldsymbol{\xi}=(\rho(1),\rho(2),\cdots,\rho(k))^T,\quad \tilde{\xi}_k=(\rho(k),\rho(k-1),\cdots\rho(1))^T,$$
$$\boldsymbol{\eta}_k=(\varphi_{k1},\varphi_{k2},\cdots,\varphi_{kk})^T,\qquad \gamma_k=(\varphi_{k1},\varphi_{k2},\cdots,\varphi_{k,\ k-1})^T$$

这时有$T^{-1}=T$，$T\boldsymbol{\rho}_kT=\boldsymbol{\rho}_k$，$T\boldsymbol{\xi}_k=\tilde{\xi}_k$，于是式(3-5-6)可写成

$$\boldsymbol{\rho}_k\boldsymbol{\eta}_k=\boldsymbol{\xi}_k \tag{3-5-9}$$

对于$k+1$则有

$$\boldsymbol{\rho}_{k+1}\boldsymbol{\eta}_{k+1}=\boldsymbol{\xi}_{k+1}$$

或

$$\begin{pmatrix}\rho_k & \tilde{\xi}_k\\ \tilde{\xi}_k^T & 1\end{pmatrix}\begin{pmatrix}\zeta_{k+1}\\ \varphi_{k+1,\ k+1}\end{pmatrix}=\begin{pmatrix}\boldsymbol{\xi}_k\\ \rho(k+1)\end{pmatrix} \tag{3-5-10}$$

可见

$$\boldsymbol{\rho}_k\zeta_{k+1}\tilde{\xi}_k\varphi_{k+1\ \ k+1}=\boldsymbol{\xi}_k$$
$$\tilde{\xi}_k^T\boldsymbol{\xi}_{k+1}+\varphi_{k+1,\ k+1}=\rho(k+1)$$

由第一式得

$$\zeta_{k+1}=\boldsymbol{\rho}_k^{-1}\boldsymbol{\xi}_k-\varphi_{k+1,\ k+1}\boldsymbol{\rho}_k^{-1}\tilde{\xi}_k=\boldsymbol{\eta}_k-\varphi_{k+1,\ k+1}T\boldsymbol{\eta}_k$$

代入第二式得

$$\begin{aligned}\rho(k+1)&=\varphi_{k+1,\ k+1}+\tilde{\xi}_k^T(\eta_k-\varphi_{k+1,\ k+1}T\boldsymbol{\eta}_k)\\ &=\varphi_{k+1,\ k+1}(1-\boldsymbol{\xi}_k^T\boldsymbol{\eta}_k)+\boldsymbol{\xi}_k^T\eta_k\end{aligned}$$

故

$$\varphi_{k+1,\ k+1}=\frac{\rho(k+1)-\tilde{\xi}_k^T\boldsymbol{\eta}_k}{1-\boldsymbol{\xi}_k^T\boldsymbol{\eta}_k} \tag{3-5-11}$$

在实际使用偏相关函数曲线估计模型的阶数时还有一个困难，这就是如何确定偏相关函数值是否真正为零。由于观察样本的长度总是有限的，偏相关的估计必然会有误差。昆诺伊尔(Qu

enouille）证明，对于AR(n)过程，当k>n时｛φ_{KK}｝近似为独立分布，且每个φ_{kk}具有零均值和近似为$1/N$的方差（N是建模所用的观察数据个数），而且当N很大时，φ_{kk}的估计值近似为正态分布。因此粗略地检验φ_{kk}是否为零，可以看φ_{kk}是否落在$\pm 2\sqrt{1}/N$范围内。

图3-5为蒸馏塔数据（见图3-2）的偏相关函数曲线。偏相关函数是调用程序PARCR计算的。

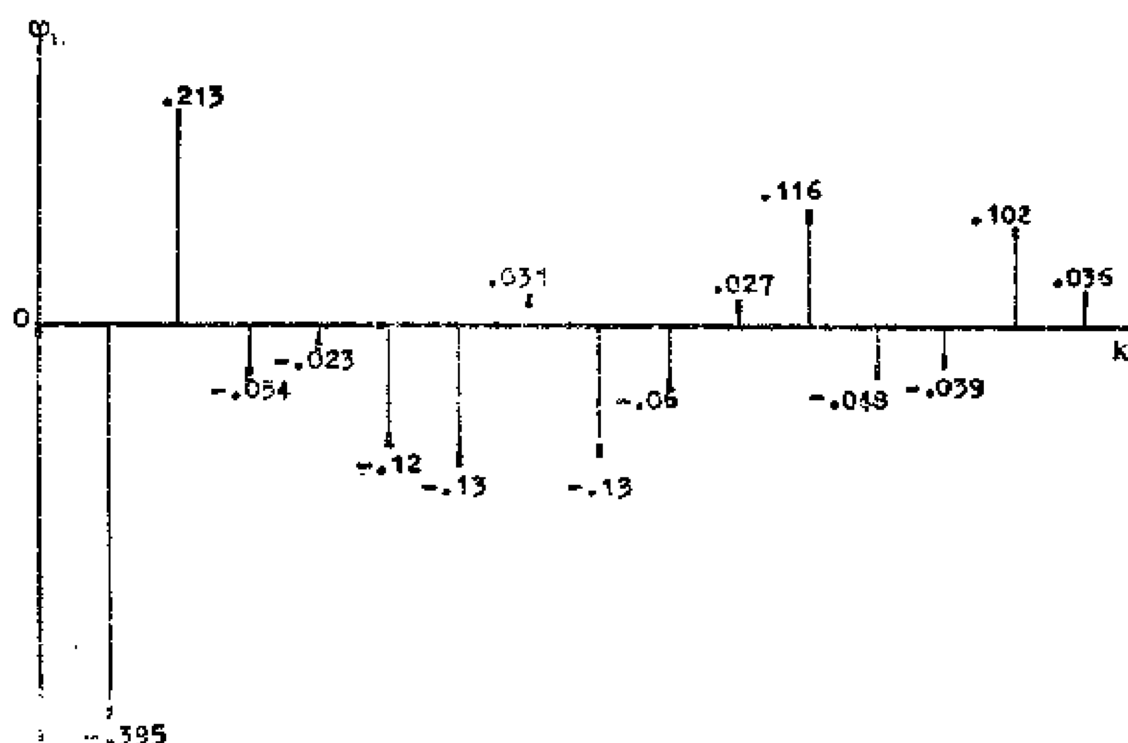


图3—5 蒸馏塔数据的偏相关函数

3. F检验用于模型定阶 在统计检验中经常用到如下两种分布：

(1) χ^2分布——如果随机变量$x_1,\cdots,x_\nu$相互独立且服从标准正态分布，即$x_t\sim\mathcal{N}(0,1)$，$t=1,2\cdots\nu$，则平方和

$$x=\sum_{t=1}^{\nu}x_t^2$$

遵从自由度为ν的χ^2分布，记作

$$x \sim \chi^2_\nu$$

其分布密度的表达式为

$$p_\nu(x) = \begin{cases} \dfrac{1}{2^{\nu/2}\Gamma(\nu/2)} \exp\left(-\dfrac{x}{2\sigma^2}\right) x^{\nu/2-1} & \text{当} x \geqslant 0 \\ 0 & \text{当} x < 0 \end{cases} \tag{3-5-12}$$

其中$\Gamma(\cdot)$是伽玛（Gamma）函数在某点的取值。

（2）F分布——如果$X \sim \chi^2_{\nu_1}$，$Y \sim \chi^2_{\nu_2}$，X与Y独立，则统计量

$$F = \frac{X}{\nu_1} \Big/ \frac{Y}{\nu_2} \tag{3-5-13}$$

的分布称为自由度为（ν_1，ν_2）的F的分布，记作$F_{\nu_1\nu_2}$或F（ν_1 ν_2），其分布密度表达式为

$$p_F(x) = \begin{cases} \dfrac{\Gamma\left(\dfrac{\nu_1+\nu_2}{2}\right)}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)} \nu_1^{\nu_1/2} \nu_1^{\nu_2/2} \dfrac{x^{\nu_1/2-1}}{(\nu_2+\nu_1 x)^{(\nu_1+\nu_2)/2}} \\ \qquad \text{当} x \geqslant 0 \\ 0 \qquad \text{当} x < 0 \end{cases} \tag{3-5-14}$$

利用F分布进行假设检验是实际中经常使用的统计检验方法。在作回归分析中，往往用F-检验来考察两个回归模型是否有显著差异，因此常被借用来判定ARMA模型的阶数。

假定我们有N个独立的随机观察值

$$\mathbf{Y} = (y_1, y_2, \cdots, y_N)^T$$

另外有r个回归因子

$$\mathbf{X}_i = (x_{i1}, x_{i2}, \cdots x_{iN})^T \qquad i = 1, 2, \cdots, r$$

我们建立回归模型

$$\mathbf{Y} = \alpha_1 \mathbf{X}_1 + \alpha_2 \mathbf{X}_2 + \cdots + \alpha_r \mathbf{X}_r + \varepsilon \tag{3-5-15}$$

其中$\varepsilon=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_N)^T$是模型残差。设$\hat{\alpha}$是模型(3-5-15)参数$\alpha=(\alpha_1,\alpha_2\cdots\alpha_r)^T$的最小二乘估计，那么，回归残差平方和为

$$Q_0=\sum_{t=1}^{N}(y_t-\hat{\alpha}_1x_{1t}-\hat{\alpha}_2x_{2t}-\cdots-\hat{\alpha}_rx_{rt})^2$$

假若我们舍弃后面s个因子，另建立一个回归模型，回归方程是

$$Y=\alpha'_1X_1+\alpha'_2X_2+\cdots+\alpha'_{r-s}X_{r-s}+\varepsilon' \qquad (3-5-16)$$

类似地，模型（3-5-16)参数α'的最小二乘估计为$\hat{\alpha}'$，相应的残差平方和是

$$Q_1=\sum_{t=1}^{N}(y_t-\hat{\alpha}'_1x_{1t}-\hat{\alpha}'_2x_{2t}-\cdots-\hat{\alpha}'_{r-s}x_{r-s,t})^2$$

我们要检验回归因子

$$X_{r-s+1},X_{r-s+2},\cdots,X_r$$

对Y的影响是否显著，即等价于检验原假设

$$H_0:\ \alpha_{r-s+1}=\alpha_{r-s+2}=\cdots=\alpha_r=0$$

是否成立。借助于F-检验我们就可以判断应该拒绝或接受原假设H_0。

关于回归分析中残差平方和的分布及假设检验有下面定理[5]

a) $Q_0\sim\sigma^2\chi^2_{N-r}$（$\sigma^2$为残差方差）

b) Q_0和Q_1-Q_0相互独立，且若原假设H_0为真，则$Q_1-Q_0\sim\sigma^2\chi^2_s$，从而

$$\frac{Q_1-Q_0}{s}\Big/\frac{Q_0}{N-r}\sim F(s,N-r)$$

有了这些结论我们可以着手检验回归因子$X_{r-s+1},X_{r-s+2},\cdots,X_r$对Y的影响是否显著，对于预先给定的显著性水平$\alpha$（一般取$\alpha$为0.05或0.01），由F分布表查出相应的$F_\alpha$值，应有下面关系式：

$$P(F \geqslant F_{\alpha}) = \alpha$$

计算统计量 $F = \frac{Q_1 - Q_0}{s} \Big/ \frac{Q_0}{N-r}$，若算得的值F大于 F_{α}，那就意味着统计量F不是$F(s, N-r)$分布（否则就意味着发生了以α为概率的小概率事件，而小概率事件在一次取样中一般被认为是不会发生的），由前述定理c)知原假设H_0为谬误，应予以拒绝，也就是说后s个回归因子X_{r-s+1}，$X_{r-s+2}, \cdots, X_r$对于随机变量Y的影响是显著的；反之，则认为不能拒绝原假设，也就是说这些因素是可以忽略的。

我们把F-检验应用于ARMA模型定阶，现在以ARMA(n, m）模型定阶为例，加以说明。我们采用过拟合的办法，先对观察数据用ARMA(n, m)模型进行拟合，再假定a_n，b_m高阶系数中某些取值为零，用F-检验准则来判定阶数降低之后的模型与ARMA(n, m)模型之间是否存在显著性差异。如果差异显著，则说明模型的阶数仍存在着升高的可能性；若差异不显著，则模型阶数可以降低。而低阶模型与高阶模型之间的差异用残差平方和来衡量。

例如，假定原假设为

H_0：$a_n = 0$，$b_m = 0$

记Q_0为ARMA(n, m)模型的残差平方和，Q_1为ARMA (n-1, m-1）模型的残差平方和，则

$$F = \frac{Q_1 - Q_0}{2} \Big/ \frac{Q_0}{N-m-n} \sim F(2, N-m-n)$$

其中N为样本长度，$r = m + n$是模型参数总个数，$s = 2$是被检验的参数个数。如果$F > F_{\alpha}$，则H_0不成立，模型阶数仍有上升的可能，否则H_0成立，即ARMA(n-1, m-1）是适合的模型。

其它类型模型的过拟合F-检验方法与上述完全类似，只需相

应改变s和r即可。

4. FPE，AIC，BIC准则

本节将要介绍另外一种最佳准则函数法，即确定出一个准则函数，该函数既要考虑用某一模型拟合时对原始数据的接近程度，同时又要考虑模型中所含待定参数的个数。建模时按照准则函数的取值确定模型的优劣，以决定取舍。使准则函数达到极小的是最佳模型。

准则函数方法首先由日本学者赤池(*Akaike*)提出，1971年他提出了一种识别*AR*模型阶数的准则，称为最小最终预报误差准则，简称为最小*FPE*（*Final Prediction Erro*）准则[17]。1973年赤池又将此方法推广到辨识ARMA模型阶数，称为最小信息准则或最小AIC（*A-Information Criterion*）准则[18]。近年来AIC准则得到广泛应用，进而又推广为BIC准则[19,20]等等。

下面对这些准则逐一加以介绍。

(1) FPE准则。FPE准则是由模型的预报误差来判明自回归模型的阶数是否恰当。这种做法在直观上是很自然的。我们用一个自回归模型去拟合某组量测数据，往往是希望借助于模型由已有的观察资料去预测未来，因此，预报效果的好坏，反过来也可以作为模型拟合优劣的检验准则。

若数据所符合的真实模型应是AR(n)，而我们用模型AR(p)($p<n$或$p>n$)去进行拟合。事实上，不论是缺参数拟合($p<n$)或是超参数($p>n$)拟合都会使预报误差的方差增大。下面我们就来说明这一点。

对于超参数拟合的情形。设$\{x_t, 1\leqslant t\leqslant N\}$所符合的的真实模型是$n$阶自回归模型，而我们用AR($p$)模型($p>n$)进行最小二乘拟合。真实模型是

$$x_t=\varphi_1 x_{t-1}+\cdots+\varphi_n x_{t-n}+\varphi_{n+1}x_{t-n-1}+\cdots+\varphi_p x_{t-p}+\varepsilon_t$$

其中 $$\varphi_{n+1}=\cdots=\varphi_{p}=0 \tag{3-5-17}$$

最小二乘拟合模型为

$$x_t=\phi_1 x_{t-1}+\cdots+\phi_p x_{t-p}+\varepsilon_t \tag{3-5-18}$$

$\widehat{\beta}=[\hat{\phi}_1,\cdots,\hat{\phi}_p]^T$是模型参数的最小二乘估计。用$\hat{x}_t^{(p)}(j)$表示t时刻用模型(3-5-18)所做的j步预报值，(见本书第六章时间序列的预报)，应有

$$\hat{x}_t^{(p)}(j)=\hat{\phi}_1 x_{t+j-1}+\cdots+\hat{\phi}_p x_{t+j-p}$$

引入符号

$$X=\begin{pmatrix} x_p & x_{p-1} & \cdots & x_1 \\ x_{p+1} & x_p & \cdots & x_2 \\ \cdots & & & \\ x_{N-1} & x_{N-2} & \cdots & x_{N-p} \end{pmatrix}=[X_1,X_2]$$

X_1是矩阵X中前n列元素组成的子矩阵，X_2是X的后$p-n$列元素组成的子矩阵，有下面关系式成立

$$X^TX=\begin{pmatrix} X_1^T \\ X_2^T \end{pmatrix}[X_1,X_2]=\begin{pmatrix} X_1^TX & X_1^TX_2 \\ X_2^TX & X_2^TX_2 \end{pmatrix}$$

令$Z=[x_{t+j-1},\cdots,x_{t+j-n},\ x_{t+j-n-1},\cdots,x_{t+j-p}]^T=[Z_1^T,Z_2^T]^T$，即对Z也做相应分解。

由最小二乘估计的无偏性（见式(3-3-6)）有

$F=\hat{x}_t^{(p)}(j)=E[\hat{\phi}_1 x_{t+j-1}+\cdots+\hat{\phi}_p x_{t+j-p}]=\phi_1 x_{t+j-1}+\cdots+\phi_p x_{t+j-p}$

现在计算预报误差的方差$E[x_{t+j}-\hat{x}_t^{(p)}(j)]^2$，注意利用$\cos[\widehat{\beta}]=\sigma_\varepsilon^2(X^TX)^{-1}$（见式(3-3-9)，其中$\widehat{\beta}=[\hat{\phi}_1,\cdots,\hat{\phi}_p]^T$）

$$\begin{aligned} E[x_{t+j}-\hat{x}_t^{(p)}(j)]^2 &= E[\sum_{i=1}^{p}(\phi_i-\hat{\phi}_i)x_{j+j-i}+\varepsilon_{t+j}]^2 \\ &= Z^T\mathrm{Cov}[\widehat{\beta}]Z+\sigma_\varepsilon^2 \\ &= \sigma_\varepsilon^2 Z^T(X^TX)^{-1}Z+\sigma_\varepsilon^2 \end{aligned}$$

将$X=[X_1, X_2]$代入，于是得

$$E[x_{t+j}-\hat{x}_t^{(p)}(j)]^2=\sigma_\varepsilon^2+\sigma_\varepsilon^2[Z_1^T, Z_2^T]\begin{bmatrix}X_1^TX_1, & X_1^TX_2\\ X_2^TX_1, & X_2^TX_2\end{bmatrix}^{-1}\begin{bmatrix}Z_1\\ Z_2\end{bmatrix} \quad (3\text{-}5\text{-}19)$$

令$\hat{x}_t^{(n)}(j)$表示只用前n个参数所得最小二乘估计相应的j步预报值，由类似推理可得

$$E[x_{t+j}-\hat{x}_t^{(n)}(j)]^2=\sigma_\varepsilon^2+\sigma_\varepsilon^2Z_1(X_1^TX_1)^{-1}Z_1 \quad (3\text{-}5\text{-}20)$$

利用分块矩阵求逆公式[5]不难证明式(3-5-19)的右端大于或等于式(3-5-20)的右端，因此得知当$p>n$时

$$E[x_{t+j}-\hat{x}_t^{(p)}(j)]^2\geqslant E[x_{t+j}-\hat{x}_t^{(n)}(j)]^2 \quad (3\text{-}5\text{-}21)$$

下面讨论缺参数拟合的情形。假定真实模型是AR(n)，即

$$x_t=\varphi_1x_{t-1}+\cdots+\varphi_nx_{t-n}+\varepsilon_t$$

或用矢量矩阵符号表示成（见式(3-3-15)）

$$Y=X\beta+\varepsilon \quad (3\text{-}5\text{-}22)$$

假定我们用AR(p)模型（$p>n$）对观察数据进行拟合，即模型为

$$x_t=\hat{u}_1x_{t-1}+\cdots+\hat{u}_px_{t-p}+\varepsilon_t' \quad (3\text{-}5\text{-}23)$$

如果用X_1表示X的前p列元素组成的子矩阵，X_2表示X的后$n-p$列元素组成的子矩阵，则$X=[X_1, X_2]$。由式（3-3-12）知模型式（3-5-23）参数的最小二乘估计为

$$\hat{u}=[\hat{u}_1,\cdots,\hat{u}_p]^T=[X_1^TX_1]^{-1}X_1^TY$$

其中$Y=[x_{p+1},\cdots,x_N]^T$

将参数向量β的前p个元素记作β_1，其余记作β_2，即

$$\beta=[\varphi_1,\cdots,\varphi_p,\varphi_{p+1},\cdots,\varphi_n]^T=[\beta_1^T, \beta_2^T]^T$$

可以计算$E\hat{u}$如下

$$E\hat{u}=(X_1^TX_1)^{-1}X_1^TEY=(X_1^TX_1)^{-1}X_1^TX\beta$$

$$=(X_1^TX_1)^{-1}X_1^T[X_1,X_2]\begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix}$$

$$=(X_1^TX_1)^{-1}X_1^TX_1\beta_1+(X_1^TX_1)^{-1}X_1^TX_2\beta_2$$

$$=\beta_1+(X_1^TX_1)^{-1}X_1^TX_2\beta_2 \tag{3-5-24}$$

再定义

$$\xi=[x_{t+j-1},\cdots,x_{t+j-n}]^T=[\xi_1^T,\ \xi_2^T]^T$$

其中 $\xi_1=[x_{t+j-1},\cdots,x_{t+j-p}]^T$

$$\xi_2=[x_{t+j-p-1},\cdots,x_{t+j-n}]^T$$

用$\hat{x}_t^{(p)}(j)$表示在时刻t用模型式(2-5-23)进行j步预报的预报值，则有

$$\hat{x}_t^{(p)}(j)=\hat{u}_1x_{t+j-1}+\cdots+\hat{u}_px_{t+j-p}=\xi_1^T\hat{u}$$

$$E[x_{t+j}-\hat{x}_t^{(p)}(j)]^2=E[\xi^T\beta-\xi_1^T\hat{u}+\varepsilon_{t+j}]^2$$

$$=\sigma_\varepsilon^2+E[\xi^T\beta-\xi_1^T\hat{u}]^2$$

$$=\sigma_\varepsilon^2+E[\xi^T\beta-\beta_1^T(\hat{u}-E\hat{u}+E\hat{u})]^2$$

$$=\sigma_\varepsilon^2+(\xi^T\beta-\xi_1^TE\hat{u})^2+E[\xi_1^T(\hat{u}-E\hat{u})]^2$$

将式(3-5-24)代入，并注意到$E[\xi_1^T(\hat{u}-E\hat{u})]^2=\sigma_\varepsilon^2\xi_1^T(X_1^TX_1)^{-1}\xi_1$

$$E[x_{t+j}-\hat{x}_t^{(p)}(j)]^2=\sigma_\varepsilon^2+[\xi_1^T\beta_1+\xi_2^T\beta_2-\xi_1^T\beta_1-\xi_1^T(X_1^TX_1)^{-1}X_1^TX_2\beta_2]^2+\sigma_\varepsilon^2\xi_1^T(X_1^TX_1)^{-1}\xi_1$$

$$=\sigma_\varepsilon^2+\sigma_\varepsilon^2\xi_1^T(X_1^TX_1)^{-1}\xi_1+[\xi_2^T\beta_2-\xi_1^T(X_1^TX_1)^{-1}X_1^TX_2\beta_2]^2 \tag{3-5-25}$$

式(3-5-25)的右端由三项组成：第一项是ε_{t+j}的方差σ_ε^2，第二项是估计$\hat{u}$值产生的预报误差方差，第三项是由于模型(3-5-23)中缺少参数β_2而产生的偏差项。当数据个数N很大，且矩阵X满足一定条件时(3-5-25)中的第二项可以是很小的量，但第三项却由于缺少参数而保持很大的量。如果我们恰用$AR(n)$模型进行拟合，那么式(3-5-25)中第三个加项不存在，而N充分大时有

$$E[x_{t+j}-\hat{x}_t^{(n)}(j)]^2\leqslant E[x_{t+j}-\hat{x}_t^{(p)}(j)]^2$$

综合上述，我们说明了不论是超参数拟合或是缺参数拟合，都会使AR模型预报误差方差增大。为使问题简化，我们考虑一步预报误差方差，要找出合理的AR模型阶数，以平衡超参数或

缺参数两种损失，这也就是寻求使一步预报误差方差达到极小的最佳AR模型阶数。

设 $\{x_t;\ 1\leqslant t\leqslant N\}$ 所适合的真实模型为$AR(n)$

$$x_t=\varphi_1 x_{t-1}+\varphi_2 x_{t-2}+\cdots+\varphi_n x_{t-n}+\varepsilon_t \qquad (3\text{-}5\text{-}26)$$

其中$E\varepsilon_t=0$，$E\varepsilon_t^2=\sigma_\varepsilon^2$，设$\varphi_i$的估计值为$\hat{\varphi}_i$ $1\leqslant i\leqslant n$，用$\hat{x}_{t-1}(1)$表示t-1时的一步预报值（即对t时刻的预报值），应有

$$\hat{x}_{t-1}(1)=\hat{\varphi} x_{t-1}+\cdots+\hat{\varphi}_n x_{t-n} \qquad (3\text{-}5\text{-}27)$$

可以证明，一步预报误差方差为

$$E[x_x-\hat{x}_{t-1}(1)]^2\approx(1+n/N)\sigma_\varepsilon^2 \qquad (3\text{-}5\text{-}28)$$

用$\hat{\sigma}_\varepsilon^2$表示以$\hat{\varphi}_i$ $1\leqslant i\leqslant n$为参数的拟合模型残差方差，可证明当N充分大时有

$$E(\hat{\sigma}_\varepsilon^2)\approx(1-n/N)\sigma_\varepsilon^2 \qquad (3\text{-}5\text{-}29)$$

换句话说，当N充分大时，$\hat{\sigma}_\varepsilon^2/(1-n/N)$是$\sigma_\varepsilon^2$的无偏估计。在式(3-5-28)中，用无偏估计代替σ_ε^2，便得到

$$E[x_t-\hat{x}_{t-1}(1)]^2\approx(1+n/N)(1-n/N)^{-1}\hat{\sigma}_\varepsilon^2 \quad (3\text{-}5\text{-}30)$$

式(3-5-28)，(3-5-29)的证明思路请参阅本书第七章7.4节。

对于以$\hat{\varphi}_i$，$l\leqslant i\leqslant n$为参数的$AR(n)$模型应有

$$\hat{\sigma}_\varepsilon^2=R(0)-\sum_{j=1}^{n}\hat{\varphi}_j R(j)$$

其中$R(0)$，$R(1)$，$\cdots R(n)$是数据序列的样本协方差函数在不同迟后时的值。

由上述可知，以$\hat{\varphi}_n,\cdots\hat{\varphi}_n$为参数的$AR(n)$模型，其最终预报误差定义为

$$\mathrm{FPE}(n)=\left(1+\frac{n}{N}\right)\left(1-\frac{n}{N}\right)^{-1}\left(R(0)-\sum_{j=1}^{n}\hat{\varphi}_j R(j)\right)$$

(3 5-31)

在解决实际问题时，我们通常是对观察数据从低阶到高阶建立

AR模型，并计算出相应的FPE值，由此确定出使最终预报误差达到极小的AR模型。

拟合的最高阶数通常取为$M(N)=\frac{1}{3}N\sim\frac{2}{3}N$，对$n=1,2,\cdots,M(N)$逐个建立AR($n$)模型并求FPE($n$)，若

$$\mathrm{FPE}(n_0)=\min_{0\leqslant n\leqslant M(N)}\mathrm{FPE}(n) \tag{3-5-32}$$

则满足(3-5-32)的AR(n_0)模型是在FPE准则之下的最佳模型。

(2) AIC准则。AIC定阶准则是赤池首先提出并成功地应用于自回归模型的分析定阶中，他将这一方法运用于水泥转窑温度的自动控制，取得了显著成效。AIC准则的理论研究近年来逐渐丰富，但仍主要限于对自回归模型定阶的理论研究。最近一些文章中，把AIC准则法引伸到回归问题定阶，并给出近似的AIC准则，使得这种定阶方法在实际中的应用得到进一步发展。在本书第八章中，我们将介绍用近似AIC准则挑选回归变元的新方法。应该指出，这些引伸的方法及近似方法，在理论研究方面还很欠缺，有待今后进一步充实，就方法本身也需要在实践中不断改进。

还应说明一点，赤池提出AIC准则时虽是针对AR模型，但这并不意味着AIC准则法具有很大的局限性。从理论上看，任何ARMA或MA模型都可以用无穷阶的AR模型逼近。因而有一种观点认为应尽可能用AR模型来描述时间序列，即使真实模型是ARMA或MA模型，也可以用高阶自回归模型来近似。因此，赤池提出的AIC准则有广泛而深刻的意义。这种方法还启发我们定义类似的准则函数，去解决其它参数估计问题。

设$\{x_t,\ 1\leqslant t\leqslant N\}$为一随机序列，我们用AR($n$)模型来描述它，$\hat{\sigma}_\varepsilon^2(n)$是拟合残差方差，我们认为它是模型阶数$n$的函数，定义AIC准则函数如下

$$\mathrm{AIC}(n)=\log\hat{\sigma}_\varepsilon^2(n)+2n/N \tag{3-5-33}$$

当阶数n增高时，式(3-5-33)中的第一项拟合残差方差$\sigma_\varepsilon^2(n)$是下降的。式中log表示取自然对数，它是单调函数，因此整个第一项是单调下降的；对给定的观察数据个数N，式(3-5-33)中的第二项随n而增长。可以想象，从$n=0$开始，逐次增加模型阶数对数据进行自回归模型拟合时，AIC(n)的值是有下降趋势的，这时起决定性作用的是模型残差方差，当达到某一阶数n_0时，AIC(n_0)值达到极小。随后，随着模型阶数继续增高，残差方差改进甚微，于是模型阶数起关键性作用，AIC(n)的值便随n而增长。对事先给定的最高阶数$M(N)$，若

$$\mathrm{AIC}(n_0)=\min_{0\leqslant n\leqslant M(N)}\mathrm{AIC}(n)$$

我们便取n_0为最佳自回归模型阶数。

将AIC准则函数用于ARMA模型定阶，方法与前述完全类似，假定我们用ARMA(n,m)模型对随机序列$\{x_t,\ 1\leqslant t\leqslant N\}$进行拟合，$\sigma_\varepsilon^2(n,m,\mu)$是残差方差，如果序列均值$\mu$也是待估计参数，那末参数个数为$n+m+1$，定义AIC函数为

$$\mathrm{AIC}(n,m,\mu)=\log\hat{\sigma}_\varepsilon^2(n,m,\mu)+2(n+m+1)/N \tag{3-5-34}$$

我们选取不同的n,m及模型参数，选取不同的均值μ，对$\{x_t\}$进行拟合，并用式(3-5-34)计算该模型相应的AIC值。然后改变模型的阶数及参数，使式(3-5-34)达到极小的模型，认为是最佳模型。

赤池的AIC准则函数方法可以引伸用于其它模型（如门限模型、混合回归模型等），解决定阶问题，我们在一些具体模型中还要做进一步说明。

(3) BIC准则及其它　在对数据序列进行模型拟合及定阶时，还可以定义与AIC函数类似的其它准则函数。BIC准则函数定义如下：

$$\mathrm{BIC}(n)=\log\hat{\sigma}_{\varepsilon}^{2}(n)+n\log N/N \tag{3-5-35}$$

若某一阶数n_0'满足

$$\mathrm{BIC}(n_0')=\min_{0\leqslant n\leqslant M(N)}\mathrm{BIC}(n)$$

其中$M(N)$含义同前，则取n_0'为最佳阶数。

与式(3-5-33)定义的AIC准则函数相比，式(3-5-35)右边第二项用$\log N$代替了系数2，一般说$\log N\gg 2$，因此AIC达到极小时所对应的阶数往往比BIC准则相应定出的阶数高，即是说一般将会有

$$n_0'\leqslant n_0$$

这说明对同一数据序列进行拟合，用AIC准则往往比用BIC准则确定的阶数高。

我们还可以定义其它类型的准则函数，如

$$\mathrm{BIC}_1(n)=\log\hat{\sigma}_{\varepsilon}^{2}(n)+Cn\log\log N/N \tag{3-5-36}$$

其中C是给定常数。定义不同的准则函数，其目的是为了对拟合残差与参数个数之间进行不同的权衡，以体现使用者对残差与阶数二者重要性的不同侧重。当然，用不同准则挑选出的最优模型，其渐近性质是不同的。例如当样本个数N趋于无穷时，用AIC准则挑选的最佳模型的阶数往往是过相容的，也就是说选定的阶数往往比真实模型的阶数高；而用BIC准则确定的最佳模型阶数往往是相容的，也就是说选定的阶数往往与真实模型的阶数相一致。

此外，在实际问题中，对不同阶数模型得到的准则函数值，往往不是理想的下凸函数，而是总的趋势符合下凸函数变化规律，同时具有随机起伏。有时可能出现准则函数值达到某值后，没有明显的增长趋势，而是随机地起伏摆动。遇到这种情形，如果适当地增大(3-5-36)式中常数系数C的值，可使准则函数值在后一段有较明显的增长趋势。

图3-6是根据蒸馏塔数据（见图3-2）拟合不同模型时得出的AIC值作出的曲线，与同样条件下的残差方差图（图3-3）比较可以看出，AIC值开始时随阶数增加而减小，而当阶数增加较多时AIC值有增长趋势，在本例中，以AR(2)（即$n=2$，$m=0$时）模型所对应的AIC值为最小，因此以二阶自回归模型来拟合是最适宜的。

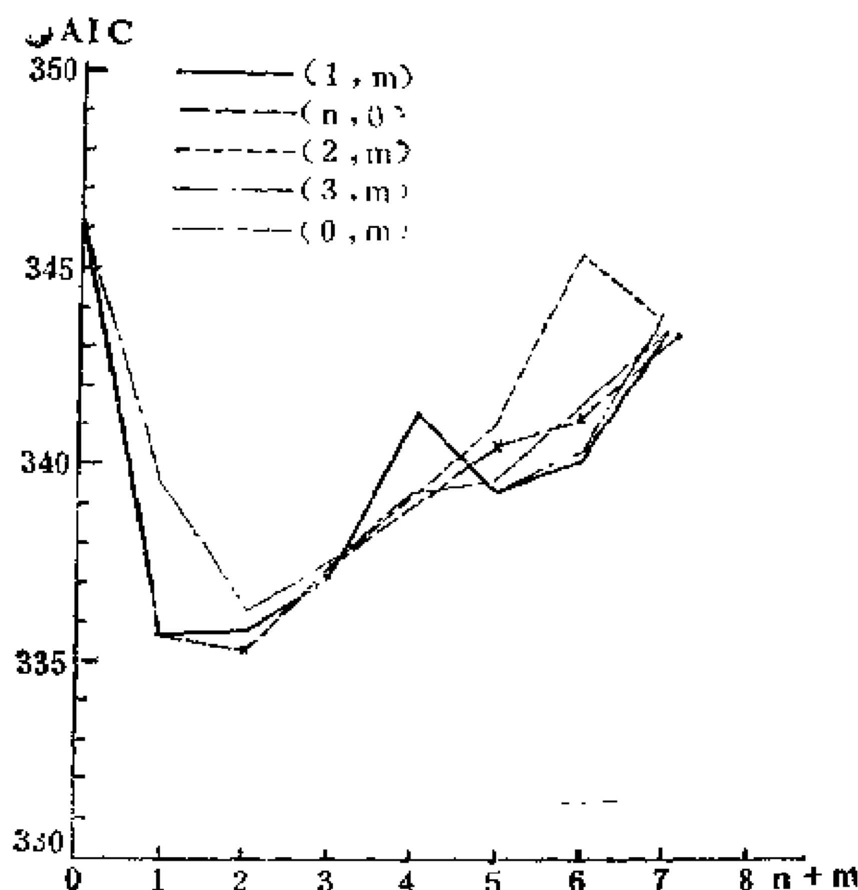


图3—6　蒸馏塔数据拟合模型AIC值

3.6　时间序列建模的基本步骤

1．波克斯—詹金斯方法

波克斯—詹金斯方法是根据样本自相关、偏相关函数的统计特性，初步判断随机序列$\{x_t\}$所适合的滑动平均、自回归阶

数。这种方法可以对AR(n)或MA(m)及低阶ARMA模型进行初步识别。

初步识别包括计算序列的均值、自相关和偏相关函数，从而确定模型的初步类型。如果样本均值显著非零，则模型表达式中应当用$\tilde{x}_t=x_t-\bar{x}$代替x_t，为检查$E[x_t]=0$是否成立，可以将样本均值$\bar{x}$和均值的标准差$S.E.[\bar{x}]$比较，而由3.1.1节知$\mathrm{Var}[\bar{x}]$在N较大时近似为$\sigma_x^2S(0)/N$，由于

$$\sigma_x^2S(0)=G(0)=\sum_{r=-\infty}^{\infty}R(r),$$

实际上用以下估计式：

$$S.E.[\bar{x}]=\sqrt{\sigma_{\bar{x}}^2}\approx\left[\frac{1}{N}\sum_{r=-(N-1)}^{N-1}\hat{R}(r)\right]^{1/2}$$

如果样本均值在$\pm2S.E.[\bar{x}]$范围内可以认为是零均值过程。

在3.5.2节中曾经指出过程的自相关函数和偏相关函数的尾部性质和过程类型的关系，而且还知道，从理论上说，MA(m)模型的自相关函数$\rho(r)$当$r>m$时截尾，AR(n)模型的偏相关函数φ_{kk}当$k>n$时截尾，考虑到它们尾部的统计性质，则它们尾部是否截止及何处截止可以根据ρ或φ_{kk}是否均落入$\pm2S.E.$范围内(例如检验$\varphi_{kk}=0$的方法是看φ_{kk}是否在$\pm2\sqrt{1/N}$范围内)。因此判断方法归结如下：

$$\varphi_{kk}\sim\mathcal{N}(0,1/N)\quad k>n \longrightarrow \mathrm{AR}(n)$$

$$\rho(r)\sim\mathcal{N}(0,1/N[1+2\sum_{i=1}^{m}\rho^2(i)])\quad r>m\longrightarrow \mathrm{MA}(m)$$

$$\varphi_{kk}\text{和}\rho(r)\text{均不截尾}\longrightarrow \mathrm{ARMA}$$

必须指出，AR过程的自相关、MA过程的偏相关以及ARMA的自相关和偏相关都不截尾，而且由于它们是齐次线性差分方程的解，因此又均受指数控制，在平稳过程情况下，尾部呈指数衰

减，这就是所谓"拖尾"。如果尾部不截止但非指数形式，则不属于上述意义下的拖尾，应当另作处理。

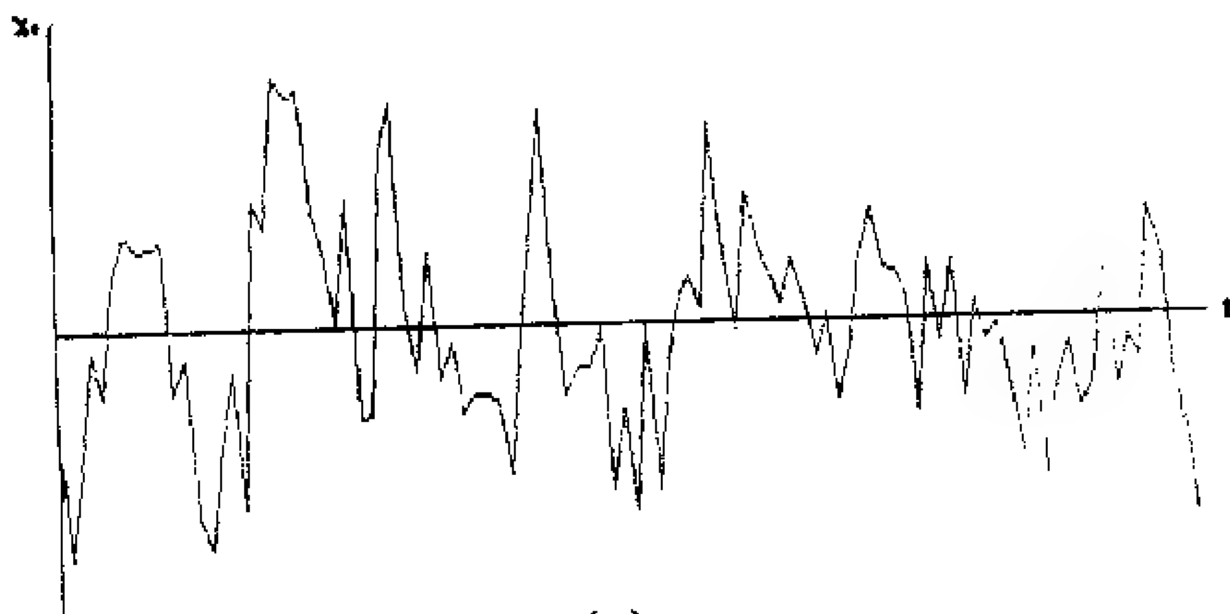
现在归纳波克斯—詹金斯的模型初步识别方法为：

(1) 若样本自相关函数$\rho(\tau)$在$\tau>m$之后截尾，则判断序列$\{x_t\}$是MA(m)序列；若样本偏相关函数φ_{kk}在$k>n$以后截尾，则判断$\{x_t\}$是AR(n)序列；若$\rho(\tau)$和φ_{kk}都不截尾，但被负指数型函数所控制（即是拖尾的），则应判断其为ARMA序列，但尚不能确定阶数。

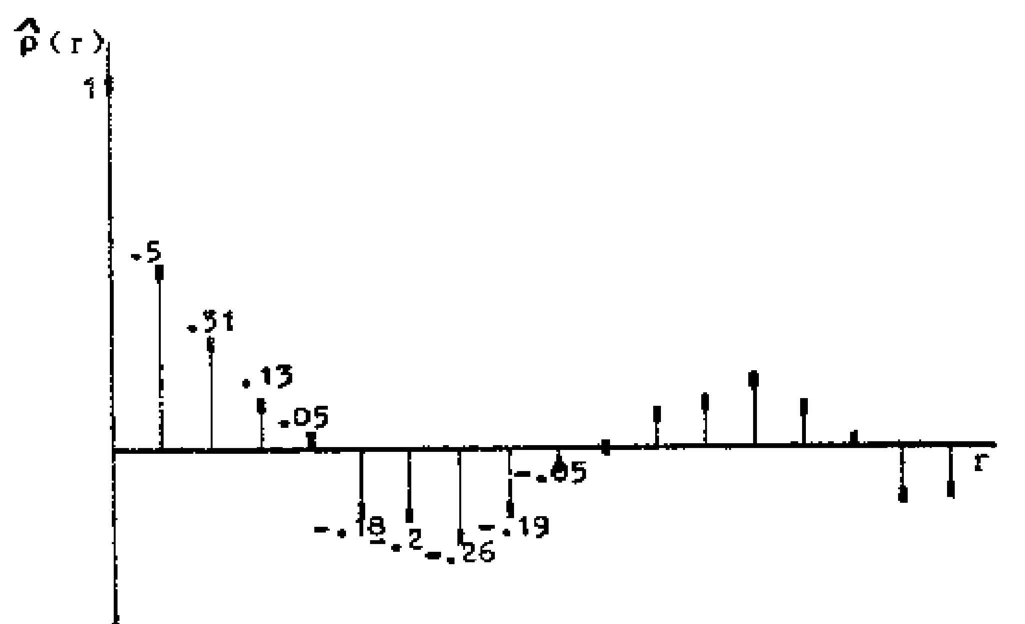
(2) 若序列$\{x_t\}$的样本自相关和偏相关函数不但都不截尾，而且（至少有一个）不是拖尾的，即下降趋势很慢，不能被负指数函数所控制，或是不具有下降趋势而是周期性变化，那么，我们便认为观察序列$\{x_t\}$具有增长趋势或季节性变化，可应用本书第八章中讲述的提取趋势性和季节性的方法，对数据进行处理。

(3) 模型的定阶。若我们已判别出序列$\{x_t\}$属于AR，MA或ARMA类型，那么就要从低阶到高阶对模型逐个进行拟合及检验。例如经鉴别判明数据序列是ARMA(n,m)类型，且n,m都不为零。我们可以从低阶到高阶逐个取(n,m)为(1,1)，(1,2)，(2,1)，(2,2)，…等值进行尝试。即对某组(n,m)所确定的ARMA(n,m)模型进行参数估计，并用节3.5中介绍的模型定阶方法，或是用F-检验决定是否接受，或是比较准则函数值决定模型的优劣。对n,m取不同值的模型进行比较，从而确定最适宜的模型。波克斯—詹金斯认为实际应用中混合模型的阶数(n,m)一般都比较低，因此逐个穷举进行模型挑选的办法是可行的。对于AR或MA模型，也是按照前述方法，从低阶到高阶逐个模型进行参数估计和模型检验。

例1. 有N=100的时间序列（见图3-7a），求得$\bar{x}=-0.08$，$\bar{x}^2=1.42$，相关函数和偏相关函数如图3-7b和c。

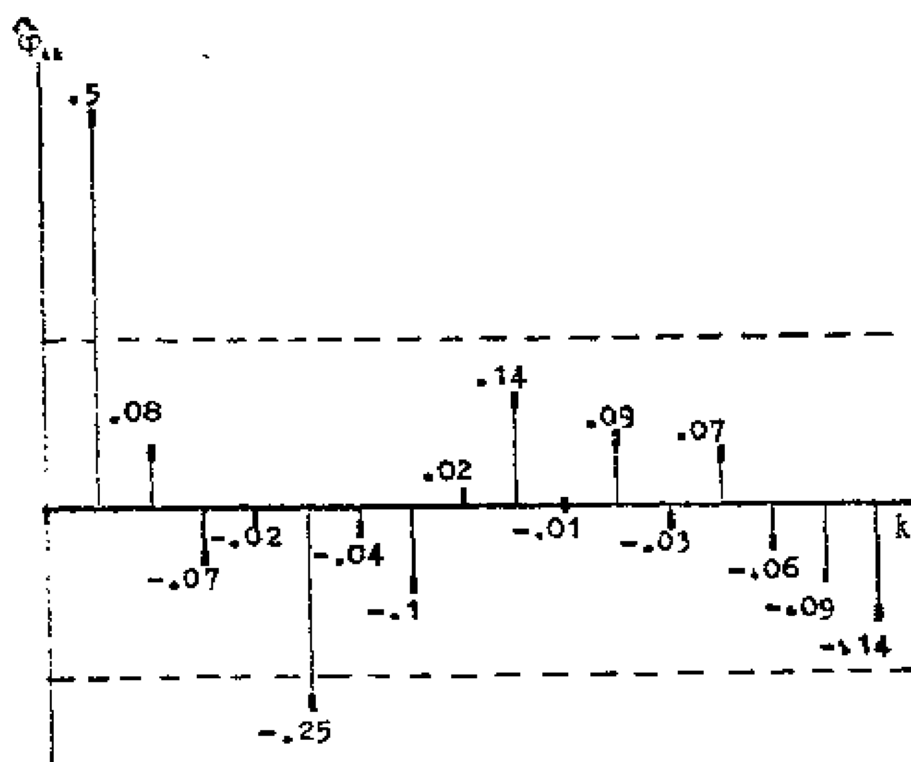


(a)



(b)

图3—7 仿真数据及其自相关和偏相关



(c)

图3－7　仿真数据及其自相关和偏相关

从$\rho(r)$看它不是MA过程，如果φ_{55}超出标准差范围可以认为是出于偶然的话，φ_{kk}是符合AR(1)模型的，否则φ_{kk}就不能和一个低阶的模型相符，也可考虑ARMA(1,1)模型的可能性，如果两者都是可接受的话，从“简约”的原则出发，先试用AR(1)模型。

首先作初始估计

AR(1)：$\hat{a}^{(0)}=-\hat{\rho}(1)=-0.499\approx-0.5$

$$\hat{\sigma}_\varepsilon^2=\hat{R}(0)[1+\sum_{i=1}^{n}\hat{a}_i\hat{\rho}(i)]=1.42(1-0.499^2)=1.07$$

$$S.E.[\bar{x}]=\sqrt{\frac{\hat{R}(0)[1+\hat{\rho}(1)]}{N[1-\hat{\rho}(1)]}}=\sqrt{\frac{1.42(1+0.499)}{100(1-0.499)}}=0.21$$

$\bar{x}$ 长显著，初始AR(1)模型取为

$$x_t - 0.50x_{t-1} = \varepsilon_t \qquad \varepsilon_t \sim \mathcal{N}(0, 1.07)$$

ARMA(1,1)：$\hat{a}^{(0)} = -\rho(2)/\rho(1) = -0.308/0.499 = -0.62$

$$\hat{b}^{(0)} = \frac{1}{2}(c \pm \sqrt{c^2 - 4})$$

其中 $c = [\hat{a}^{(0)2} - 2\rho(2) + 1]/[\rho(1) + \hat{a}^{(0)}]$

$$= (0.6172^2 - 2 \times 0.308 + 1)/(0.499 - 0.6172) = -6.472$$

$$\hat{b}^{(0)} = \frac{1}{2}(-6.472 + \sqrt{6.472^2 - 4}) = -0.16$$

$$\hat{\sigma}_\varepsilon^2 = \hat{R}(0)(1 - \hat{a}^2)/(1 - 2\hat{a}\hat{b} + \hat{b}^2)$$

$$= 1.42(1 - 0.6172^2)/(1 - 2 \times 0.6172 \times 0.1585 + 0.1585^2)$$

$$= 0.83$$

$$S.E.[\bar{x}] = \sqrt{\frac{\hat{R}(0)}{N}\left[1 + \frac{2\rho^2(1)}{\rho(1) - \rho(2)}\right]}$$

$$= \sqrt{\frac{1.42}{100}\left(1 + \frac{2 \times 0.499^2}{0.499 - 0.308}\right)} = 0.23$$

故 $\bar{x}$ 不显著，ARMA(1,1) 的初始模型取为

$$x_t - 0.62x_{t-1} = \varepsilon_t - 0.16\varepsilon_{t-1} \quad \varepsilon_t \sim \mathcal{N}(0, 0.83)$$

在精估计后可得二模型的参数如下：

AR(1)：　$\hat{a} = -0.516$　$S.E.[\hat{a}] = 0.087$　$\hat{\sigma}_\varepsilon^2 = 1.046$

ARMA(1,1)：　$\hat{a} = -0.605$　$S.E.[\hat{a}] = 0.147$

$\hat{b} = -0.120$　$S.E.[\hat{b}] = 0.181$　$\hat{\sigma}_\varepsilon^2 = 1.04$

从两个模型 $\hat{\sigma}_\varepsilon^2$ 之比可见 $\frac{1.046}{1.04} = 1.005 < F_{98, 97; 0.05}$ $= 1.25$，所以没有实质性的差别。另外，从ARMA参数看，$\hat{b}$ 不显著（和 $S.E.[\hat{b}]$ 比较），$\hat{a}$ 也和AR模型参数无明显差别，因此没有显著的理由要用ARMA(1,1) 而不用较简约的AR(1)模型，事实上，如果在 $(1 - 0.51B)x_t = \varepsilon_t$ 两边乘以 $(1 - 0.12B)$ 则得

$(1-0.63B+0.06B^2)x_t=(1-0.12B)\varepsilon_t$，这很接近ARMA (1，1)的参数。这种多余附加的参数（称参数冗余）反映在ARMA参数的$S.E.$值比AR要大得多。

例2，对$N=256$序列计算得$\bar{x}=-0.34$，$s^2=1.71$，其自相关和偏相关分别示于图3.8a)和b)。

从$\rho(r)$看很可能是MA(1)过程，ϕ_{kk}也符合，但根据$Var[\rho(r)]$的巴特利公式，当$r>1$时$S.E.[\rho(r)]=0.076$，而在$r>1$之后只有$\rho(2)$略超出$2S.E.=0.15$。所以可能是MA(2)或ARMA(1,1)，若按MA(1)考虑，由

$$S.E.[\bar{x}]=\sqrt{\frac{\hat{R}(0)[1+2\hat{\rho}(1)]}{N}}=\left[\frac{1.71}{256}(1+2\times0.393)\right]^{1/2}$$

$$=0.109$$

可见$\bar{x}$明显非零，参数的初始估计为

$$\hat{b}^{(0)}=2\rho(1)/[1+\sqrt{1-4\rho^2(1)}]$$
$$=2\times0.393/(1+\sqrt{1-4\times0.393^2})=0.485$$

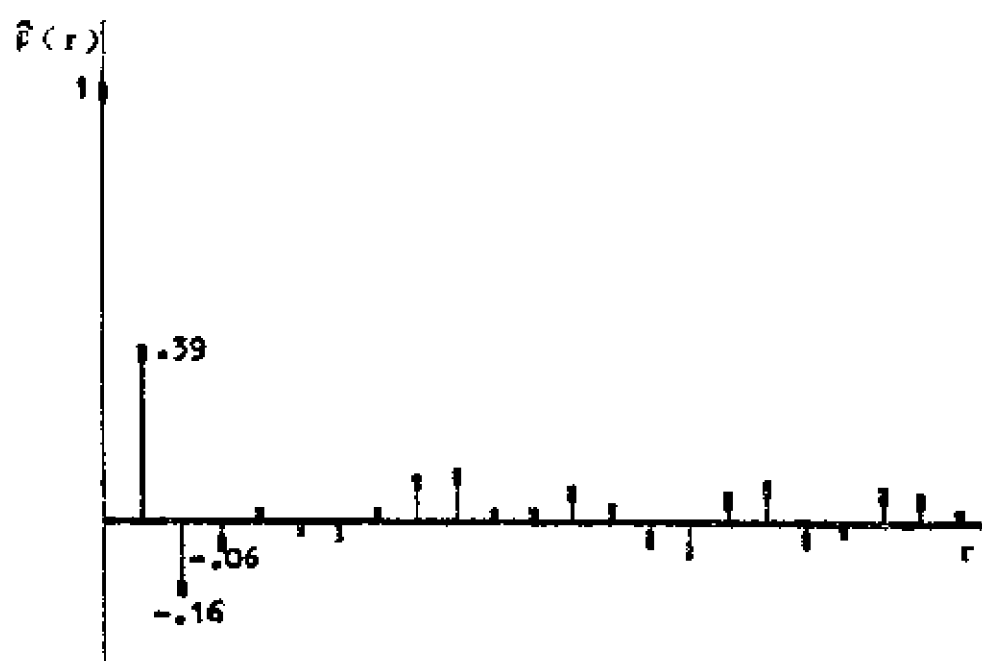
$$\hat{\sigma}_\varepsilon^2=\hat{R}(0)/(1+\sum_{i=1}^{1}\hat{b}_i^2)=1.71/(1+0.485^2)=1.384$$

于是初始模型为

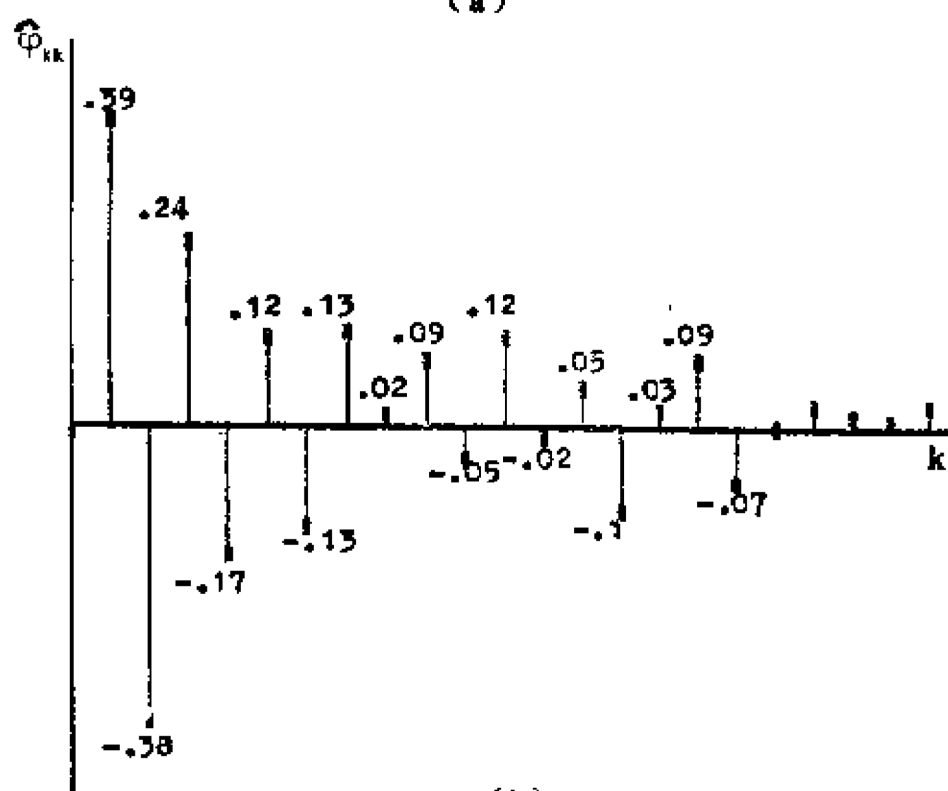
$$x_t+0.34=\varepsilon_t+0.485\varepsilon_{t-1}\qquad \varepsilon_t\sim\mathcal{N}(0,1.384)$$

由于有可能是ARMA(1,1)，因此如果在对MA(1)作精估计的同时也作ARMA(1,1)和MA(2)的估计，可得出以下结果

模型	参数估计	估计的$S.E.$	$\hat{\sigma}_\varepsilon^2$
MA(1)	$\hat{b}=0.830$	0.035	0.833
ARMA(1,1)	$\hat{a}=-0.102$	0.073	1.08
	$\hat{b}=0.858$	0.0373	
MA(2)	$\hat{b}_1=0.4921$	0.058	0.866
	$\hat{b}_2=-0.3128$	0.0585	



（a）



（b）

图3—8 自相关、偏相关图形

从以上结果看，可以不考虑ARMA(1,1)和MA(2)。

2. 潘迪特—吴（Pandit—Wu）方法的建模 吴贤铭和潘迪特在波克斯—詹金斯方法的基础上，经过实践和进一步发展，在1977年提出了系统建模的新方法，称之为动态数据系统（DDS—Dynamic Data System)建模方法。该方法从应用角度对波克斯—詹金斯方法做了一定程度的改进，在机械制造、造纸、原子能、化工及其它领域都有成功的应用。

在建模策略上，潘迪特—吴采用ARMA(n,n-1)模型逼近序列$\{x_t\}$，这是由于DDS方法最初是在研究连续系统和离散系统的对应关系时提出的。在文献〔6〕中指出，线性定常连续系统在白噪声作用下的反应，经过均匀间隔离散采样所得到的时间序列，总可以用离散的ARMA(n,n-1) 模型表示。依据这一点，对ARMA(n,m) 模型的确定可以先沿着直线方向进行（见图3—9）。亦即把二维(n,m) 搜索化为一维搜索问题，令n由小到大增加，每改变一次阶数（$n,n-1$），可根根F-检验判断阶数增大是否必要。

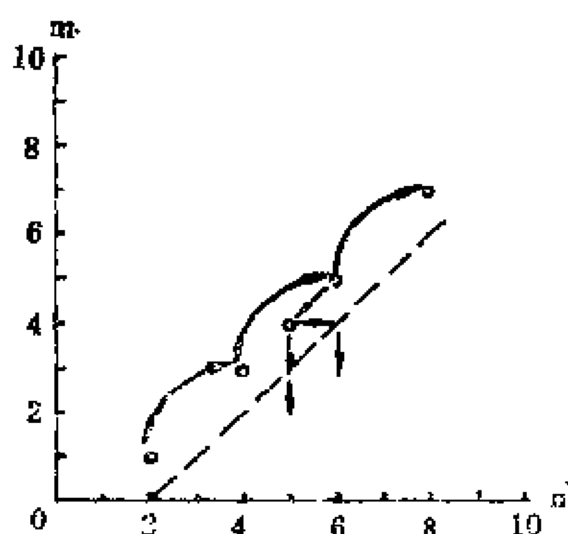


图3—9 定ARMA(n, n-1) 逼近ARMA模型

关于n的取值方式，潘迪特—吴建议从$n=2$开始，按ARMA（$2n,2n-1$）的方式进行，即ARMA（2,1），ARMA(4,3)，ARMA(6,5)…等等。其理由是：(1)阶数增加2相当于差分方程特征方程的根增加2，这样做可以避免强使模型增加一个实根而造成失误。而当根数增加二时，它既可以是共轭复根，也可以是两个实根（当实际为一个实根时，拟合结果其中将有一个实根接近于零），因此比较合理。(2)实际物理系统的自由度随着它的复杂程度而增加，而自由度每增加一，阶数便增加二。

当F检验表明升阶并不对残差平方和的减小有显著改善时，这并不意味着原来的低阶模型ARMA($2n,2n-1$)就是最简约的，还可以检查其系数a_{2n}及b_{2n-1}的置信区间是否包含零，如果是则进一步分析AR阶数或MA阶数是否有降低的可能。图3-10为上述定阶搜索方案的流程图，根据该图可以制定自动定阶的计算机程序。

潘迪特—吴认为在拟合模型之前，可以不作差分或季节性算子的预处理（见第八章8.1节），而从拟合过程和结果中识别序列中的趋势和周期性。如果有必要时可以用确定性模型和随机性模型的组合来描述一个完整的过程。

具体的做法是先用一个具有确定性函数的模型，如

$$y_t=\sum_{i=1}^{u} R_i e^{\gamma_i t}+\sum_{j=1}^{v} B_j e^{\gamma_j t}\sin(j\omega t+\psi_j)+x_t$$

该式右边x_t是ARMA过程，而其余各项是参数待定的指数或具有指数振幅的正弦函数。在用最小二乘方法定出y_t确定性部分的参数后，若剩余误差已接近平稳，则将它用ARMA模型拟合。最后再把确定性部分和随机性ARMA模型叠加在一起，用该组合模型对数据序列进行拟合。

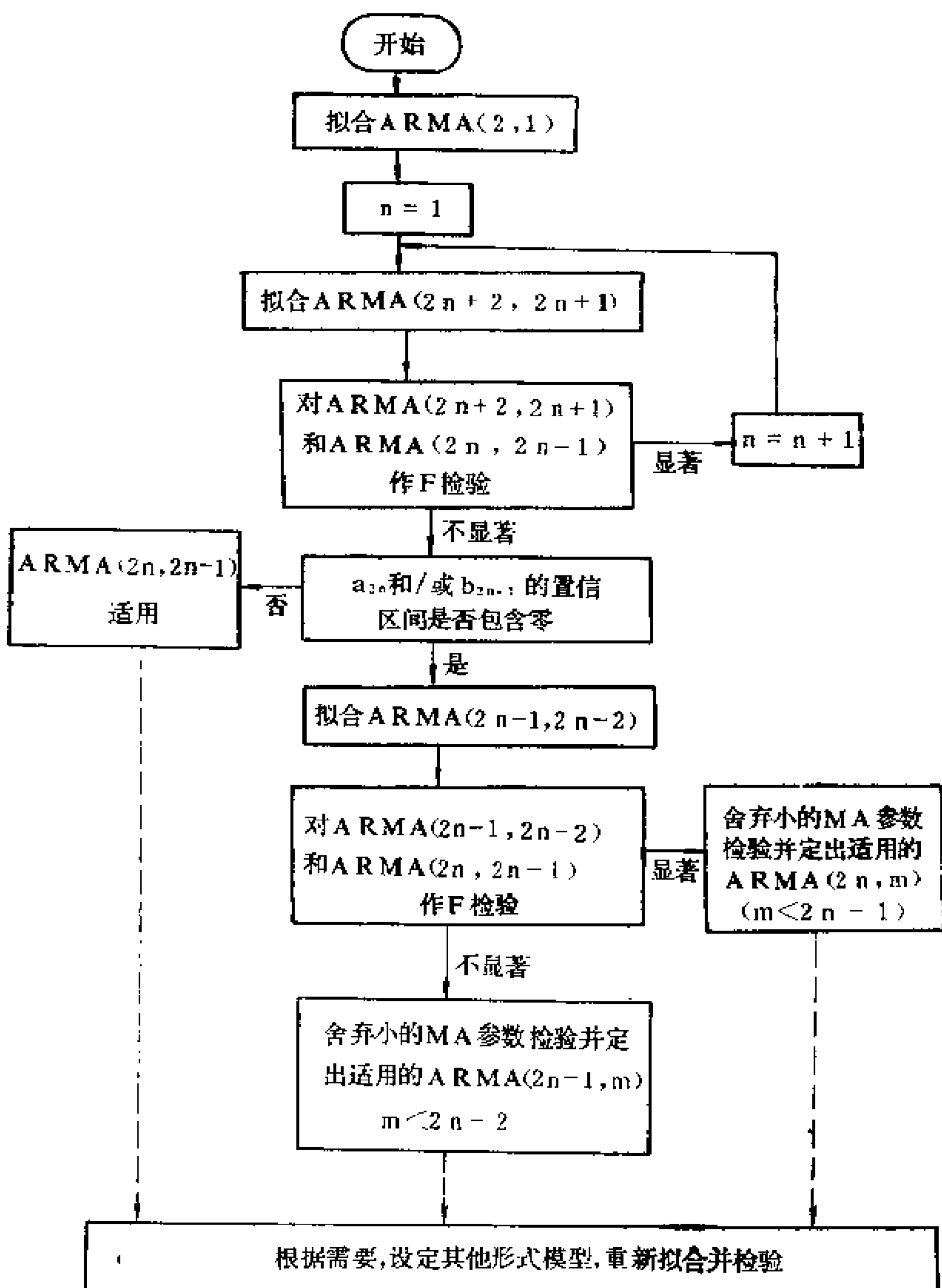


图3—10 潘迪特—吴建模方法流程图

3．长自回归白噪化建模方法 在前面我们曾提到，单纯自回归模型的定阶、参数估计远比ARMA模型简单，因此不少学者提倡用高阶自回归（或称长自回归）模型来近似ARMA序列。在〔8〕中并对这一思想的可靠性进行了理论证明，指出若自回归阶数n_N与样本个数N有如下关系

$$n_N\sqrt{\log N/N}\to 0 \qquad (N\to\infty)$$

则AR(n_N)的谱能相当好地一致逼近序列的真实谱。因此，用长自回归模型逼近ARMA序列时，阶数n_N应随样本增加而增高，需要估计的参数也随之增多。为了充分利用长自回归逼近的可靠性，并适当控制被估计参数个数。近年来有人提出了一种“长自回归白噪化”ARMA模型建模方法。该方法用足够高的自回归模型拟合随机序列，用求解线性方程组的办法代替非线性最小二乘求极值问题，从而避免了迭代计算不收敛的麻烦，并大大减少了计算量。该方法的中心思想是承认用长自回归模型逼近原序列的可靠性，因而对平稳零均值观察数据序列$\{x_t,\ 1\leqslant t\leqslant N\}$取$n_N$为充分大的数所得到的AR($n_N$)模型能足够好地拟合$\{x_t\}$。换句话说，由长自回归模型AR($n_N$)得出的残差序列$e_{n_N}$，$e_{n_{N+1}}$，…$e_N$可以近似地认为是白噪声序列，并且可以把$e_{n_N}$，$e_{n_{N+1}}$，…$e_N$作为已知值使用，结合原数据序列$\{x_t\}$一起进行ARMA模型定阶及参数估计。

长自回归白噪化建模方法具体步骤如下：

（1）建立长自回归模型AR(n_N)。阶数n_N取为

$$n_N=(\log N)^{1+\delta}$$

选δ为一正数：$0\leqslant\delta\leqslant 1$，AR($n_N$)的自回归参数为

$$\hat{\alpha}=(\hat{\alpha}_1,\ \hat{\alpha}_2,\ \cdots\hat{\alpha}_{n_N})^T$$

由耶尔—瓦克尔方程得到

$$\hat{\alpha}=\hat{R}_{n_N}^{-1}\hat{\eta}_{n_N}$$

其中

$$\widehat{R}_{n_N}=\begin{pmatrix}\hat{R}(0)\cdots\hat{R}(n_N-1)\\ \cdots \\ \hat{R}(n_N-1)\cdots\hat{R}(0)\end{pmatrix}$$

$$\widehat{\eta}_{n_N}=(-\hat{R}(1),\ -\hat{R}(2),\ \cdots,\ -\hat{R}(n_N))^T$$

模型白噪声ε_t的方差估计值为

$$\hat{\sigma}_{n_N}^2=\hat{R}(0)+\widehat{\eta}_N^T\widehat{R}_{n_N}^{-1}\widehat{\eta}_{n_N}$$

(2) 求长自回归模型残差，并检验其独立性。用(1)中所得$\widehat{\alpha}$及样本值$x_1,\cdots,x_N$计算残差

$$\hat{\varepsilon}_t=x_t+\sum_{j=1}^{n_N}\hat{\alpha}_j x_{t-j},\qquad t=n_N+1,\cdots,N$$

检验$\{\hat{\varepsilon}_t\}$是否是独立白噪声（见第一章1.4节），若不独立，则增大n_N，重新进行(1)(2)；若$\{\hat{\varepsilon}_t\}$是独立白噪声序列则进行下一步。

(3) 对给定阶数n,m，由数据序列$\{x_t\}$及(2)中得到的残差序列$\{\hat{\varepsilon}_t\ \ n_N+1\leqslant t\leqslant N\}$，求ARMA$(n,\ m)$模型参数的最小平方和估计。记模型参数$\beta^T=(a^T,\ b^T)$，其中$a^T=(a_1,\cdots a_n)$，$b^T=(b_1,\cdots,b_m)$，平方和函数为

$$Q(\widehat{\beta})=\sum_{t=n_N+m}^{N}(x_t+a_1x_{t-1}+\cdots+a_nx_{t-n}-b_1\hat{\varepsilon}_{t-1}-\cdots-b_m\hat{\varepsilon}_{t-m})^2$$

极小化$Q(\widehat{\beta})$可得

$$\begin{cases}\dfrac{\partial Q(\widehat{\beta})}{\partial a_j}=2\displaystyle\sum_{t=n_N+m}^{N}(x_t+a_1x_{t-1}+\cdots+a_nx_{t-n}\\ \qquad -b_1\hat{\varepsilon}_{t-1}-\cdots-b_m\hat{\varepsilon}_{t-m})x_{t-j}=0\quad j=1,2,\cdots n\\ \dfrac{\partial Q(\widehat{\beta})}{\partial b_i}=2\displaystyle\sum_{t=n_N+m}^{N}-(x_t+a_1x_{t-1}+\cdots+a_nx_{t-n}-\\ \qquad b_1\hat{\varepsilon}_{t-1}-\cdots-b_m\hat{\varepsilon}_{t-m})\hat{\varepsilon}_{t-i}=0\quad i=1,2,\cdots m\end{cases}$$

当N很大时，上面方程组可用下式近似

$$\hat{\beta}=\begin{pmatrix}\hat{a}\\ \hat{b}\end{pmatrix}=\begin{pmatrix}\hat{R}^{x}_{nn} & -\hat{R}^{\varepsilon x}_{nm}\\ -\hat{R}^{x\varepsilon}_{mn} & \hat{R}^{\varepsilon}_{mm}\end{pmatrix}^{-1}\begin{pmatrix}\hat{\eta}^{x}_{n}\\ \hat{\eta}^{x\varepsilon}_{m}\end{pmatrix}$$

ARMA(n,m)的残差方差估计为

$$\sigma^2_{nm}=\hat{R}(0)-\begin{pmatrix}\hat{a}\\ \hat{b}\end{pmatrix}^T\begin{pmatrix}\hat{\eta}^{x}_{n}\\ \hat{\eta}^{x\varepsilon}_{m}\end{pmatrix}$$

式中 $\hat{R}^{x}_{nn}=[\hat{R}(j-k)] \quad 1\leqslant j,\ k\leqslant n$

$$\hat{\eta}^{x}_{n}=[-\hat{R}(1),-\hat{R}(2),\cdots,-\hat{R}(n)]^T,$$

$$\hat{R}(l)=\frac{1}{N-n_N}\sum_{t=n_N+1}^{N-1}x_{t+l}x_t$$

$$\hat{R}^{\varepsilon}_{mm}=[\hat{R}^{\varepsilon}(j-k)] \quad 1\leqslant j,\ k\leqslant m$$

$$\hat{R}^{\varepsilon}(l)=\frac{1}{N-n_N}\sum_{t=n_N+1}^{N-1}x_t\varepsilon_{t-l}\cong\hat{R}^{\varepsilon}(-l)$$

$$\hat{R}^{x\varepsilon}_{nm}=[\hat{R}^{x\varepsilon}(j-k)] \quad l\leqslant j\leqslant n,\ l\leqslant k\leqslant m,$$

$$\hat{R}^{x\varepsilon}(l)=\frac{1}{N-n_N}\sum_{t=n_N+1}^{N=1}x_t\varepsilon_{t-l}\cong\hat{R}^{\varepsilon x}(-l)$$

$$\hat{\eta}^{x\varepsilon}_{m}=[\hat{R}^{x\varepsilon}(1),\ \hat{R}^{x\varepsilon}(2),\cdots,\hat{R}^{x\varepsilon}(m)]^T$$

$$\hat{R}^{\varepsilon x}_{mn}=[\hat{R}^{\varepsilon x}(j-k)] \quad 1\leqslant j\leqslant m,\ 1\leqslant k\leqslant n$$

$$\hat{R}^{\varepsilon x}(l)=\frac{1}{N-n_N}\sum_{t=n_N+1}^{N-1}\varepsilon_{t-l}x_{t=l}\cong\hat{R}^{x\varepsilon}(-l)$$

(4) 在适当阶数范围内（例如$0\leqslant n\leqslant n_N$，$0\leqslant m\leqslant m_N$，$n_N$，$m_N$为事先选定值），对$n,m$重复进行(1)—(3)，并按照某种定阶准则，确定最佳模型阶数及参数。注意，在进行模型挑选时，可利用消去变换（见第八章8.2节）以减少求解线性方程组的次数。例如设n为任意固定的自回归阶数，只需利用消去变换的结果就可以比较ARMA$(n,0)$，ARMA$(n,1)$，…，ARMA(n,m_N)

各模型的残差平方和。

4. 参数初始估计的逆函数法 在ARMA模型的参数估计中，通常要用递推方法解非线性问题，而开始递推的参数初值总希望其尽可能接近最终解，同时又不难获得。确定初始估计的方法很多，如节3.2中介绍的矩估计方法。这些方法的共同特点是通过解线性方程得出模型的参数初值，因为线性方程组求解本身无需初值，而其结果可作为非线性递推计算的出发点。

对于$\{x_t\}$的均值，我们显然可以用样本均值作为初估计。自回归和滑动平均参数的初值可以用ARMA序列的矩估计方法，而利用本节介绍的逆函数方法是比矩估计更为有效的方法。

我们知道，如果过程是可逆的，即

$(1+a_1B+a_2B^2+\cdots+a_nB^n)\ x_t=(1+b_1B+b_2B^2+\cdots+b_mB^m)\ \varepsilon_t$ 可用等价的逆转形式——无限阶的AR模型表示

$$(1-I_1B-I_2B^2-\cdots-)\ x_t=\varepsilon_t$$

其中I_i和a_i，b_i的关系为（设$n>m$）

$$
\begin{aligned}
a_1&=b_1-I_1\\
a_2&=b_2-b_1I_1-I_2\\
a_3&=b_3-b_1I_2-b_2I_1-I_3\\
&\vdots\\
a_i&=b_i-b_1I_{i-1}-b_2I_{i-2}-\cdots-b_{i-1}I_1-I_i\\
&\vdots\\
a_n&=-b_1I_{n-1}-b_2I_{n-2}-\cdots-b_mI_{n-m}-I_n\\
0&=-b_1I_n-b_2I_{n-1}-\cdots-b_mI_{n+1-m}-I_{n+1}\\
&\vdots
\end{aligned}
$$

不难看出，在$n>m$情况下，当$i\geqslant n+1$时有

$$(1+b_1B+b_2B^2+\cdots+b_mB^m)\ I_i=0$$

如果$m>n$，则当$i>m+1$时上式成立，故有

$$(1+b_1B+b_2B^2+\cdots+b_mB^m)\ I_i=0\quad i>\max(n,m)$$

这是一个线性方程组，ARMA模型的逆函数I_t有无限多项，如果取$i=\max(n,m)+1$到$i=\max(n,m)+m$的（共m个）I_i值，则可列出m个线性方程，由此解出$b_1,\cdots,b_m$，将这些b_i和$i\leqslant\max(n,m)$的I_i值代入线性方程组

$$a_i=b_i-b_1I_{i-1}-\cdots-b_{i-1}I_1-I_i$$

便可得a_i。

由上述可知，首先要求出逆函数值$I_i,i=1,2,\cdots,\max(n,m)+m$。考虑到在过程为可逆的情况下$|I_i|\to 0$（$i\to\infty$）。因此可以取足够大的$p$值使下式近似成立

$$(1-I_1B-I_2B^2-\cdots-I_pB^p)\ x_t\cong\varepsilon_t$$

显然我们可以拟合一个p阶AR模型，取$p\geqslant\max(n,m)+m$

$$(1+\tilde{a}_1B+\tilde{a}_2B^2+\cdots+\tilde{a}_pB^p)x_t=\varepsilon_t$$

其中参数$\tilde{a}_i$可根据观察数据用线性方法估计得到。令

$$I_i=-\tilde{a}_i \qquad i=1,2,\cdots,p$$

便可得所需的逆函数值。若所取的p大于$\max(n,m)+m$，即求解b_i的方程式个数多于m时，可以用线性最小二乘方法求解b_i。

但要注意，这里求出的b_i并不一定保证可逆性条件的满足，因为I_i毕竟是由有限阶的AR过程近似无限阶的AR而得出的。为了判断可逆与否，可作求根分解

$$1+b_1B+\cdots+b_mB^m=(1-\gamma_1B)(1-\gamma_2B)\cdots(1-\gamma_mB)$$

检查$|\gamma_i|$ $(i=1,2,\cdots,m)$是否小于1。若某个γ值的模大于1，则取其倒数值$1/\gamma$代替γ。例如将因子$(1-2B)$用$(1-0.5B)$代替。在完成所有必要的替换后，根据新的因子乘积得出满足可逆条件的新的b_i。可以证明，这样做并不影响原模型的相关函数或谱密度，即不改变平稳过程的统计性质。

关于根的求倒和相应的参数计算可用附录二中的程序INGUS。

第四章　周期图与加窗谱估计

在第二章2.4节给出了平稳过程的频率域函数—功率谱、积分功率谱和谱展式。这一章则讨论如何根据给定的时间序列在频率域内实现关于周期和谱特性的估计，对单变量平稳过程来说主要就是指功率谱密度函数，特别是当过程只包含纯粹连续的谱时，其统计特性用功率谱密度比积分频率更为方便。

和时间域中的离散采样一样，频率域函数的实际计算往往也要离散化，因此离散富氏变换(DFT)及其快速算法（即快速富氏变换FFT）成为不可缺少的工具，我们把采样定理，DFT，FFT的性质原理及应用放在附录中介绍。

4.1　隐周期的估计

我们先介绍如何在带有纯粹随机噪声（如量测误差等）的情况下判断周期信号的存在并估计其频率和幅值，这对分析序列中的隐含周期分量是很有用的，同时它也体现了频域方法的基本思想。这里假设离散过程的数学模型为

$$x_t = \sum_{i=1}^{K} c_i \cos(2\pi f_i t + \varphi_i) \varepsilon_t \tag{4-1-1}$$

式中K，$\{c_i\}$，$\{f_i\}$ ($i=1, 2, \cdots, K$)为常数，$\{\varphi_i\}$是在$(-\pi, \pi)$内均匀分布的独立随机变量，$\{\varepsilon_t\}$是独立于$\{\varphi_i\}$的纯随机过程，$E(\varepsilon_t)=0$，$E(\varepsilon_t^2)=\sigma_\varepsilon^2$(未知)。

1. **周期图分析**　面临的问题是：根据已有的N个观察x_1，x_2，…，x_N，如何估计式(4-1-1)中的一组未知参数：

$$\{K, c_i(i=1, \cdots, K), f_i(i=1, \cdots K), \sigma_\varepsilon^2\}$$

式(4-1-1)亦可写成

$$x_t = \sum_{i=1}^{K}(a_i\cos 2\pi f_i t + b_i \sin 2\pi f_i t) + \varepsilon_t \qquad (4\text{-}1\text{-}2)$$

其中$a_i = c_i\cos\varphi_i$，$b_i = -c_i\sin\varphi_i$（即$c_i = \sqrt{a_i^2+b_i^2}$，$\varphi_i = \tan^{-1}(-b_i/a_i)$）

如果已有关于频率｛f_i｝的先验知识，K自然亦属已知，式(4-1-2)可以看作是一个多变量线性回归模型，以a_i和b_i为未知参数，可用最小二乘法将

$$Q = \sum_{t=1}^{N}\left[x_t - \sum_{i=1}^{K}(a_i\cos 2\pi f_i t + b_i \sin 2\pi f_i t)\right]^2 \qquad (4\text{-}1\text{-}3)$$

极小化来估计相应的a_i和b_i。但实际上｛f_i｝及周期分量个数K通常是未知的，因此在尚未确定$\{f_i\}$之前还不能对式(4-1-2)模型作最小二乘估计。

在伴有噪声ε_t的情况下用"周期图"方法搜索隐含周期$\{f_i\}$的基本思想是：假定f_1估计为$\hat{f}_1$，则可由$\hat{f}_1$求估值$\hat{a}_1$和$\hat{b}_1$(如用最小二乘估计)，如果估计的$\hat{f}_1$恰好就是或者很接近f_1，那末$\hat{a}_1$和$\hat{b}_1$就近乎a_1和b_1，$(\hat{a}_1^2+\hat{b}_1^2)$显然非零。反之，如果$\hat{f}_1$和$f_1$相差甚远或者说模型中出现的是别的频率，则所估计的系数是模型中实际不存在的，那末$(\hat{a}_1^2+\hat{b}_1^2)$必近于零。用一组足够密的频率$\hat{f}_1$，$\hat{f}_2$，$\hat{f}_3$，…作试探并绘出模的平方$(\hat{a}_k^2+\hat{b}_k^2)$对$\hat{f}_k$的图形，则当$\hat{f}_k$和$\{f_i\}$中的某一值接近时$(\hat{a}_k^2+\hat{b}_k^2)$将异于零，否则就近于零。根据平方模的值足够大处所对应的f便可判断其频率，并进而估计相应的振幅。

为了导出计算$\hat{a}_i$和$\hat{b}_i$的公式，先讨论给定频率时a_i和b_i的最小二乘估计。将式(4-1-3)对a_i和b_i求导并令其为零得正规方程

$$\sum_{t=1}^{N} x_t\cos 2\pi f_j t = \sum_{i=1}^{K}\hat{a}_i\alpha_{ij} + \sum_{i=1}^{K}\hat{b}_i\beta_{ij} \quad j=1,\cdots,k \quad (4\text{-}1\text{-}4)$$

$$\sum_{t=1}^{N} x_t \sin 2\pi f_j t = \sum_{i=1}^{K} \hat{a}_i \beta_{ij} + \sum_{i=1}^{K} \hat{b}_i \gamma_{ij} \quad j=1,\cdots,k \quad (4\text{-}1\text{-}4)$$

其中

$$\alpha_{ij} = \sum_{t=1}^{N} \cos 2\pi f_i t \cos 2\pi f_j t$$

$$\beta_{ij} = \sum_{t=1}^{N} \sin 2\pi f_i t \cos 2\pi f_j t \qquad (4\text{-}1\text{-}5)$$

$$\gamma_{ij} = \sum_{t=1}^{N} \sin 2\pi f_i t \sin 2\pi f_j t$$

当样本长度N（这里采样间隔$\Delta=1$，故$N\Delta=N$）是各个周期（$1/f_i$）的整数倍p_i时，式(4-1-4)解的形式是很简单的，因各周期分量($i=1,\cdots,K$)的频率为

$$f_i = p_i/N \qquad 0 \leqslant p_i \leqslant \lfloor N/2 \rfloor \qquad (4\text{-}1\text{-}6)$$

利用正余弦函数的正交关系：

$$\sum_{t=1}^{N} \cos\left(\frac{2\pi pt}{N}\right) \cos\left(\frac{2\pi qt}{N}\right) = \begin{cases} 0 & 0 \leqslant p \neq q \leqslant \lfloor N/2 \rfloor \\ N/2 & 0 < p = q < N/2 \\ N & p = q = 0 \text{或} N/2 (N\text{偶}) \end{cases}$$

$$\sum_{t=0}^{N} \sin\left(\frac{2\pi pt}{N}\right) \cos\left(\frac{2\pi qt}{N}\right) = 0 \quad \text{所有} p,q \qquad (4\text{-}1\text{-}7)$$

$$\sum_{t=1}^{N} \sin\left(\frac{2\pi pt}{N}\right) \sin\left(\frac{2\pi qt}{N}\right) = \begin{cases} 0 & 0 \leqslant p \neq q \leqslant \lfloor N/2 \rfloor \\ N/2 & 0 < p = q < N/2 \\ 0 & p = q = 0 \text{或} N/2 (N\text{偶}) \end{cases}$$

可见当f_i为式(4-1-6)，且没有p_i为零或$N/2$时

$$\begin{aligned} &\alpha_{ij} = \gamma_{ij} = 0 && \text{所有} i \neq j \\ &\beta_{ij} = 0 && \text{所有} i, j \qquad (4\text{-}1\text{-}8) \\ &\alpha_{ii} = \gamma_{ii} = N/2 && \text{所有} i \end{aligned}$$

代入式(4-1-4)得

$$\hat{a} = \frac{2}{N}\sum_{t=1}^{N} x_t \cos 2\pi f_i t$$
$$\hat{b}_i = \frac{2}{N}\sum_{t=1}^{N} x_t \sin 2\pi f_i t \qquad (4\text{-}1\text{-}9)$$

下面来考察这种估计的偏差性质，由均值

$$E[\hat{a}_i] = \frac{2}{N}\sum_{t=1}^{N} E[x_t]\cos 2\pi f_i t \qquad (4\text{-}1\text{-}10)$$

因$E[\varepsilon_t]=0$，有$E[x_t]=\sum_{i=1}^{K}(a_i\cos 2\pi f_i t + b_i \sin 2\pi f_i t)$，代入式(4-1-10)并利用正交关系式(4-1-8)可得

$$E[\hat{a}_i] = a_i \quad (\text{所有}i) \qquad (4\text{-}1\text{-}11)$$

类似有$E[\hat{b}_i]=b_i$(所有i)，这表明估计是无偏的。此外，由式(4-1-9)可得

$$\mathrm{Var}[\hat{a}_i] = \left(\frac{2}{N}\right)^2 \sigma_\varepsilon^2 \sum_{t=1}^{N}\cos^2 2\pi f_i t = 2\sigma_\varepsilon^2/N \qquad (4\text{-}1\text{-}12)$$

类似有$\mathrm{Var}[\hat{b}_i]=2\sigma_\varepsilon^2/N$　由式(4-1-8)可求得

$$\mathrm{Cov}[\hat{a}_i, \hat{b}_i] = 0 \quad (\text{所有}i)$$

σ_ε^2的估计值为

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{N-2K}\sum_{t=1}^{N}\left[x_t - \sum_{i=1}^{K}(\hat{a}_i\cos 2\pi f_i t + \hat{b}_i \sin 2\pi f_i t)\right]^2 \qquad (4\text{-}1\text{-}13)$$

有了特定频率下$\hat{a}_i$，$\hat{b}_i$的估计算式，下面不难引出周期图的定义：

N个观察序列x_1，…，x_N的周期图为

$$\hat{I}_N(f) \equiv \frac{2}{N}\left|\sum_{t=1}^{N} x_t e^{-j2\pi f t}\right|^2 \quad -\frac{1}{2} \leqslant f \leqslant \frac{1}{2} \qquad (4\text{-}1\text{-}14)$$

它也可写成

$$\tilde{I}_N(f)=[A(f)]^2+[B(f)]^2 \qquad -\frac{1}{2}\leqslant f\leqslant\frac{1}{2} \qquad (4\text{-}1\text{-}15)$$

$$A(f)\equiv\sqrt{\frac{2}{N}}\sum_{t=1}^{N}x_t\cos 2\pi ft$$

$$B(f)\equiv\sqrt{\frac{2}{N}}\sum_{t=1}^{N}x_t\sin 2\pi ft$$

实际计算中$\tilde{I}_N(f)$的频率取值是离散的—0, 1/N, 2/N, …, 故记

$$\tilde{I}_p\equiv\tilde{I}_N(f_p) \qquad f_p=p/N,\ p=0,1,\cdots,\lfloor N/2\rfloor$$

$$\tilde{I}_p=[A(f_p)]^2+[B(f_p)]^2=\frac{N}{2}[\hat{a}_i^2+\hat{b}_i^2] \qquad (4\text{-}1\text{-}16)$$

由于$E[\hat{a}_i^2]=(E[\hat{a}_i])^2+\mathrm{Var}[\hat{a}_i]\simeq a_i^2+2\sigma_e^2/N$,同样，$E[\hat{b}_i^2]\simeq b_i^2-2\sigma_e^2/N$，因此

$$E[\tilde{I}_p]=\frac{N}{2}[E(\hat{a}_i^2]+E[\hat{b}_i^2]) \simeq \frac{N}{2}(a_i^2+b_i^2)+2\sigma_e^2$$

$$=\frac{N}{2}c_i^2+2\sigma_e^2 \qquad (4\text{-}1\text{-}17)$$

当f_p的值等于x_t中所含周期分量的频率f_i时，$E[\tilde{I}_p]=O(N)$，周期图有大的峰值，而当$|f_p-f_i|\gg 0$时 $E[\tilde{I}_p]\simeq 2\sigma_e^2$，即$E[\tilde{I}_p]=O(1)$ 由于在$\tilde{I}_N$定义式(4-1-14)中引入系数$2/N$，使峰值更加明显。

以上关于$E[\tilde{I}_p]$的考虑只是当$f_i=p_i/N(i=1,\cdots,K)$时才是准确的，否则是近似的。要想知道f_p接近但不等于f_i以及f_i不限为$1/N$的倍数时周期图是否还有明确的峰值，需要对周期图的统计特性作进一步的讨论。

2. 周期图的样本统计特性 (1)先讨论x_t为纯粹随机过程

时$\tilde{I}_p$的样本分布，即$x_t\equiv\varepsilon_t$且$c_i=0$（所有的i，即$K=0$），又假定$\{x_t\}$为正态过程。这里得出的结果将在周期图的峰值检验中用到。

由于$\{x_t\}$是独立正态随机变量构成的序列（设均值为0，方差σ_x^2），作为它的线性组合，$A(f_p)$也是零均值正态分布，其方差为（利用正交关系式(4-1-7)）

$$\mathrm{Var}[A(f_p)]=\frac{2\sigma_x^2}{N}\left(\sum_{t=1}^{N}\cos^2 2\pi f_p t\right)$$

$$=\begin{cases}\sigma_x^2 & f_p=p/N,\ p\neq 0, N/2(N\text{偶})\\ 2\sigma_x^2 & f_p=0\end{cases} \tag{4-1-18}$$

类似地，$B(f_p)$也是零均值正态分布，方差为

$$\mathrm{Var}[B(f_p)]=\begin{cases}\sigma_x^2 & p\neq 0,\ N/2(N\text{偶})\\ 0 & p=0\text{或}N/2\end{cases} \tag{4-1-19}$$

还可证明

$$\mathrm{Cov}[A(f_p),\ B(f_p)]=\frac{2\sigma_x^2}{N}\left(\sum_{t=1}^{N}\cos^2 2\pi f_p t\sin 2\pi f_q t\right)$$

$$=0\qquad\text{所有}p,q \tag{4-1-20}$$

$$\mathrm{Cov}[A(f_p),\ A(f_q)]=\mathrm{Cov}[B(f_p),\ B(f_q)]=0\quad p\neq q$$

由于$\{A(f_p)\}$和$\{B(f_p)\}$均为正态而相互无关，故为独立随机变量，它们的平方和服从自由度为2的χ^2分布，即

$$\tilde{I}_p=[A(f_p)]^2+[B(f_p)]^2=\sigma_x^2\chi_2^2\qquad p\neq 0,\ N/2 \tag{4-1-21}$$

而当$p=0$或$N/2$时

$$\tilde{I}_p=[A(f_p)]^2=(2\sigma_x^2)\chi_1^2\qquad p=0,\ N/2 \tag{4-1-22}$$

考虑到χ_ν^2分布的均值和方差分别为ν和2ν，有

$$E[\tilde{I}_p]=2\sigma_x^2\qquad\text{所有}p \tag{4-1-23}$$

$$\operatorname{Var}[\tilde{I}_p]=\begin{cases}4\sigma_x^4 & p\neq 0,\ N/2\\ 8\sigma_x^4 & p=0,\ N/2\end{cases} \tag{4-1-24}$$

（2）现在来看当$c_i\neq 0$时$E[\tilde{I}_N(f)]$的一般表达式，根据周期图的定义式(4-1-14)可改写成

$$\tilde{I}_N(f)=\frac{2}{N}\left[\sum_{t=1}^{N}x_t e^{-j2\pi ft}\right]\left[\sum_{s=1}^{N}x_s e^{j2\pi fs}\right]$$

$$=\frac{2}{N}\left[\sum_{t=1}^{N}\sum_{s=1}^{N}x_t x_s\cos(t-s)2\pi f\right]$$

令$r=t-s$，并利用自协方差的估计式$\hat{R}(r)$可得

$$\tilde{I}_N(f)=2\sum_{r=-(N-1)}^{N-1}\left[\frac{1}{N}\sum_{t=1}^{N-|r|}x_t x_{t+|r|}\right]\cos 2\pi rf$$

$$=2\sum_{r=-(N-1)}^{N-1}\hat{R}(r)\cos 2\pi rf \tag{4-1-25}$$

由式(3-1-6)有

$$E[\tilde{I}_N(f)]=2\sum_{r=-(N-1)}^{N-1}\left(1-\frac{|r|}{N}\right)R(r)\cos 2\pi rf \tag{4-1-26}$$

由于模型式(4-1-1)中的$\{\varepsilon_t\}$和周期分量彼此独立，$\{\varepsilon_t\}$的自协方差函数为$\sigma_\varepsilon^2\delta_{0,r}$($r=0$时$\delta_{0,r}=1$，其他$r$时$\delta_{0,r}=0$)，周期分量的自协方差为$\frac{1}{2}\sum_{i=1}^{K}c_i^2\cos 2\pi f_i r$，因此$\{x_t\}$的自协方差函数为

$$R(r)=\frac{1}{2}\sum_{i=1}^{K}c_i^2\cos 2\pi f_i r+\sigma_\varepsilon^2\delta_{0,r}$$

将它代入式(4-1-26)并经化简可得

$$E[\tilde{I}_N(f)]=2\sigma_\varepsilon^2+\frac{1}{2}\sum_{i=1}^{K}c_i^2[2\pi F_N(f+f_i)$$
$$+2\pi F_N(f-f_i)] \tag{4-1-27}$$

其中

$$2\pi F_N(\theta)=\frac{1}{N}\frac{\sin^2(\pi\theta N)}{\sin^2(\pi\theta)} \tag{4-1-28}$$

称$F_N(\theta)$为费泽(*Fejer*)核，其曲线如图4-1所示。令$\varphi=\pi\theta N$，式(4-1-28)可表为$2\pi F_N(\theta)\simeq N\ (\sin^2\varphi)/\varphi^2\to N$（当$\theta\to 0$），因此$F_N(\theta)$在原点处峰值为$N/(2\pi)$。$F_N(\theta)$随$\theta$增大而减小，小峰值近似出现在$\theta=\pm\frac{3/2}{N}$，$\pm\frac{5/2}{N}$，$\pm\frac{7/2}{N}$，…，而在$\pm 1/N$，$\pm 2/N$，$\pm 3/N$处$F_N(\theta)=0$。

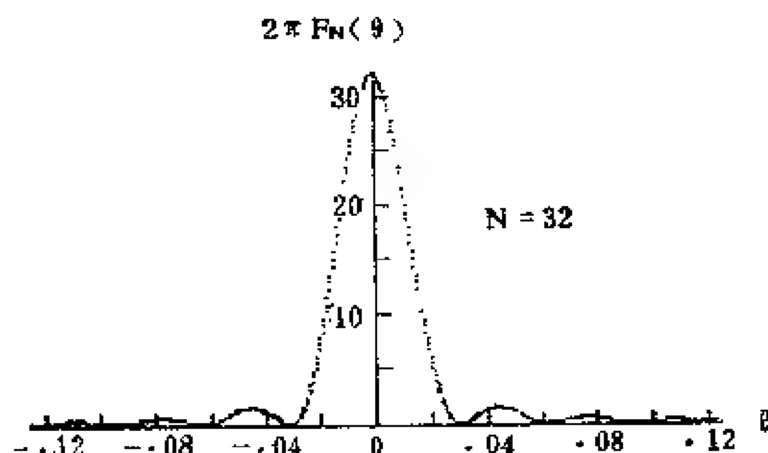


图4-1 费泽核（N=32）

整个$E[\bar{I}_N(f)]$图形由常数项$2\sigma_\varepsilon^2$和以$(f\pm f_i)$，$i=1,\cdots K$为中心的2π个费泽核所组成(见图4-2)。由于$|\theta|\gg 1/N$时$F_N(\theta)=O(1/N)$，因此若各周期分量的频率相距足够远(如$|f_i\pm f_j|\gg 1/N$，$i\neq j$)，则在$f=f_i$附近的周期图除常数项外，主要由式(4-1-27)中和该f_i所对应的项—$2\pi[F_N(f+f_i)+F_N(f-f_i)]$所决定。当$f$恰为$f_i$时，$E[\bar{I}_N(f)]$的峰值就是$2\sigma_\varepsilon^2+\frac{c_i^2}{2}N$（和式（4-1-17）一

致)。计算中f按p/N（$p=0, 1, \cdots, \lfloor N/2 \rfloor$）取值，为了不致漏过峰值点，间隔要取得足够细密，例如可取$p/(2N)$或$p/(4N)$等。

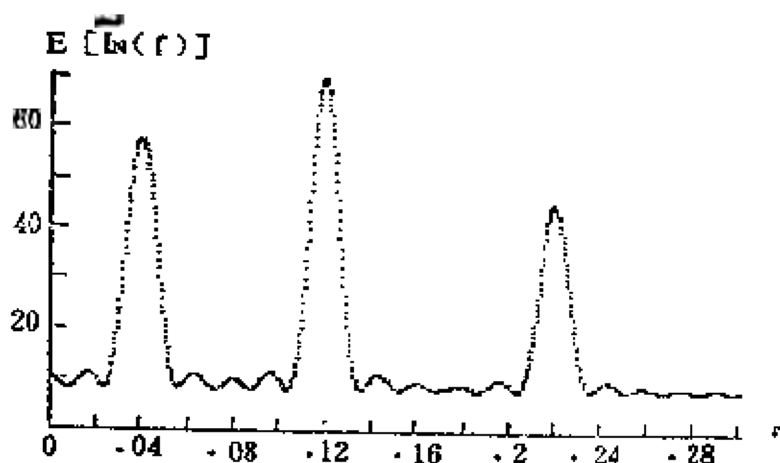


图4-2　由三个费泽核组成的E〔$\tilde{I}_N$（f）〕示例

3.　**周期图的峰值检验**　虽然一般总可以从周期图上找到一些峰值，但还不能立即断定哪些对应着x_t的真正周期分量，因为甚至$c_i=0$(所有i)的情况下，也会因随机采样引起周期图纵坐标的波动而出现虚假的峰值，为此需用统计检验方法来确定所出现的峰值是否比$c_i=0$时的情况要显著得多。

通常按标准频率$f_p=p/N(p=0, 1, \cdots, \lfloor N/2 \rfloor)$画出周期图，然后先检验其中的最大峰值。对模型式(4-1-1)所取的零假设为

H_0：$c_i=0$　　所有i　(或等效为$K=0$)

由式(4-1-21)知$p=1, \cdots, \lfloor N/2 \rfloor$时($\tilde{I}_p/\sigma_x^2$)的概率密度为$p(x)=e^{-x/2}/2(0\leqslant x<\infty)$，对任何$z$有

$$P[(\tilde{I}_p/\sigma_x^2)\leqslant z]=\int_0^z e^{-x/2}/2dx=1-e^{-z/2} \qquad (4\text{-}1\text{-}29)$$

以γ表最高峰所对应的($\tilde{I}_p/\sigma_x^2$)，即

$$\gamma=\frac{1}{\sigma_x^2}\left[\max_{1\leqslant p\leqslant\lfloor N/2\rfloor}(\tilde{I}_p)\right] \tag{4-1-30}$$

利用式(4-1-29)并考虑到$\{\tilde{I}_p\}$的独立性质可得

$$P[\gamma>z]=1-P[\gamma\leqslant z]=1-P[(\tilde{I}_p/\sigma_x^2)\leqslant z,\ 所有p]$$
$$=1-(1-e^{-z/2})^{\lfloor N/2\rfloor} \tag{4-1-31}$$

这就是H_0假设下γ的概率分布。假定事先已知σ_x^2，则可用单侧检验判断γ是否大于某一z_a值(z_a由显著水平选定)。实际上要用由数据求得的σ_x^2估值代入式(4-1-30)，考虑到在H_0假设下$E[\tilde{I}_p]=2\sigma_x^2$(所有$p\geq 1$)，故$E\left[\sum_{p=1}^{\lfloor N/2\rfloor}\tilde{I}_p\right]=2\lfloor N/2\rfloor\sigma_x^2$，因此$\sigma_x^2$的无偏估计为$\sum_{p=1}^{\lfloor N/2\rfloor}\tilde{I}_p/(2\lfloor N/2\rfloor)$，从而得统计量

$$g^*=\frac{\max(\tilde{I}_p)}{\frac{1}{2\lfloor N/2\rfloor}\sum_{p=1}^{\lfloor N/2\rfloor}\tilde{I}_p} \tag{4-1-32}$$

由于N很大时上式分母可认为是常值σ_x^2，故g^*和γ的分布相同，利用式(4-1-31)

$$P[g^*>z]\simeq 1-(1-e^{-z/2})^{\lfloor N/2\rfloor} \tag{4-1-33}$$

可对$\max(\tilde{I}_p)$作大样本检验，根据选择的显著水平α，由$P[g^*>z_a]=\alpha$定出z_a，若计算的g^*超过z_a，则$\max(\tilde{I}_p)$在$100\alpha\%$水平上是显著的，从而拒绝H_0，即$\{x_t\}$含有周期分量。

费歇(*Fisher*)提出的另一种$\max(\tilde{I}_p)$检验方法是根据

$$g=\max(\tilde{I}_p)\Big/\sum_{p=1}^{\lfloor N/2\rfloor}\tilde{I}_p \tag{4-1-34}$$

称其为费歇统计量，这里是考虑到$\tilde{I}_p$和$\Sigma\tilde{I}_p$都含有比例因子σ_x^2，所以其比值的分布不依赖于σ_x^2。费歇证明了在H_0假设下g的分布

为

$$P[g<z]=n(1-z)^{n-1}-\frac{n(n-1)}{2}(1-2z)^{n-1}+\cdots\cdots+$$

$$(-1)^{a}\frac{n!}{a!(n-a)!}(1-az)^{n-1} \quad (4\text{-}1\text{-}35)$$

其中：$n=\lfloor N/2\rfloor$，a为小于$1/z$的最大整数。

由α选定z_α使得$P[g>z_\alpha]=\alpha$，然后根据z_α进行检验。

费歇检验只涉及最大的周期图峰值，惠多(*Whittle*)提出将检验推广到第二大峰值。设$\tilde{I}_{p1}$是经检验为显著的最大峰值，则第二大峰值($\tilde{I}_{p2}$)检验的统计量取为

$$g'=\frac{\tilde{I}_{p2}}{\left[\sum_{p=1}^{[N/2]}\tilde{I}_p\right]-\tilde{I}_{p1}} \quad (4\text{-}1\text{-}36)$$

g'的分布和式(4-1-35)类似，但以$(n-1)$代替n。如果检验表明$\tilde{I}_{p2}$是显著的，则继续对第三大峰值作检验，依此类推可得K个显著的峰值，这就是模型式(4-1-1)中周期分量的个数。

格伦南德(*Grenander*)和罗森布来特(*Rosenblatt*)导出了在零假设H_0(x_t为正态纯随机过程)情况下，周期图第r个峰值(以$\tilde{I}_{(r)}$表示)和周期图纵坐标总和之比(记作$g_{(r)}=\tilde{I}_{(r)}/\Sigma_p\tilde{I}_p$)的概率分布为

$$P[g_{(r)}>z]=\frac{n!}{(r-1)!}\sum_{j=1}^{n}\frac{(-1)^{j-r}(1-jz)^{n-1}}{j(n-j)!(j-r)!} \quad (4\text{-}1\text{-}37)$$

利用这个分布，可以检验在给定r时H_0假设是否成立（即x_t中是否正好有r个周期项）。

在附录三表5中给出了在显著水平为0.01和0.05情况下不同n，r参数时的g值。由于g值变化很平滑，因此如果需要查n，r处于表中已给参数之间时的g值可通过内插来求取。在实际应用

中，由于r未知，可以这样来使用该表：先取$r=1$时的g，接受所有振幅大于该值者，如果从第j个振幅起($j>1$)小于该值，则取$r=j$时的g值，接受比这个新值大的那些振幅，再做下去就是取振幅不满足该值要求的新的r值，并按此值得新的g值，这样一直做到新的g值比被检验的z值大为止，所有那些较小的振幅在这个显著水平上都被拒绝了。

下面是一个齿轮双面啮合综合测量的例子，两个啮合齿轮的中心距数据如图4.3所示(a)图为前512点曲线，b)图为相连的后512点曲线)。可以看到一个缓慢的周期变化，而由100点曲线的图形（图4.4）则可看出高频变化的成分。利用隐周期分析检验

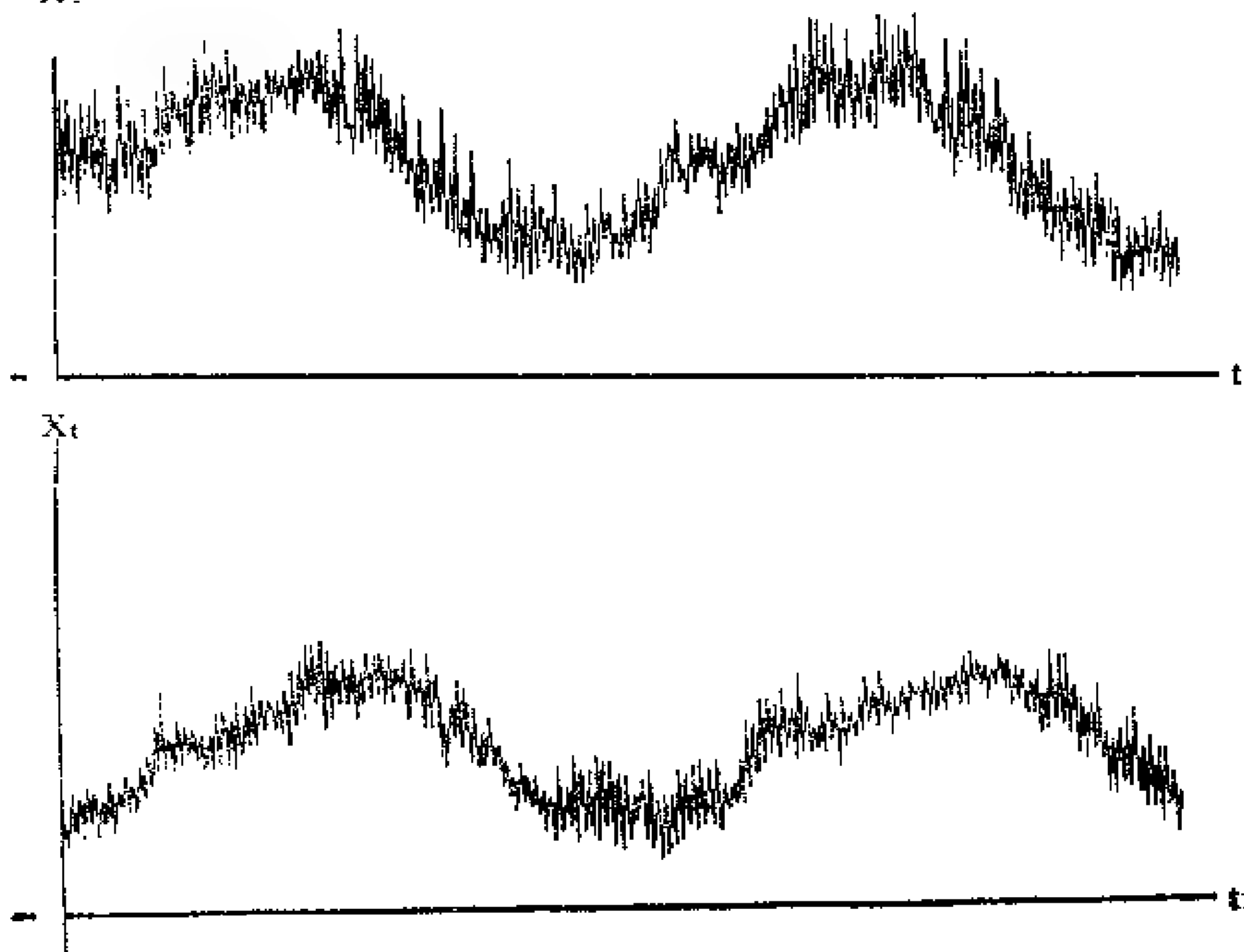


图4-3　两个啮合齿轮的中心距实测数据（1024点）

程序（TESTP）和检验表对上述1024个数据进行计算分析，可得按大小顺序排列的g值（按式(4-1-34)计算）如表4-1

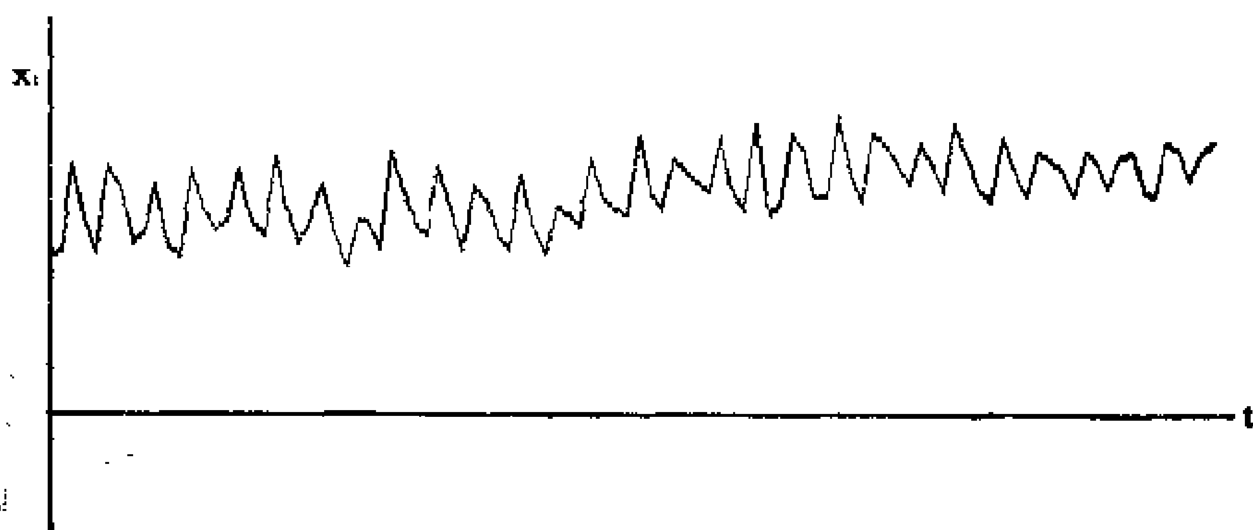


图4-4　图4-3数据的局部放大（100点）

表4-1

频率	0.003906	0.002929	0.001953	0.29
g值	0.60155	0.18539	0.05229	0.02693

这些周期分量与测量齿轮和被测量齿圈的齿周径向跳动以及齿形误差、基节偏差等具体因素有关。通过分析峰值频率的位置，将可探明齿轮啮合误差的组成成分，为查找误差源提供信息。

4.2　功率谱密度的周期图估计

1．修正周期图与功率谱估计　现在考虑当观察到的$\{x_t\}$是具有连续谱函数的随机过程，对所有的f,其功率谱密度存在。下面将利用周期图的思想对功率谱密度作出估计。

根据x_t的谱展式(2-4-72)有

$$\sum_{t=1}^{N} x_t e^{-j2\pi f t} = \sum_{t=1}^{N}\left[\int_{-1/2}^{1/2} e^{j2\pi\theta t}\, dZ(\theta)\right] e^{-j2\pi f t}$$

$$= \int_{-1/2}^{1/2}\left(\sum_{t=1}^{N} e^{-j2\pi(\theta-f)t}\right) dZ(\theta)$$

而

$$\sum_{t=1}^{N} e^{j2\pi\phi t} = e^{j(N+1)\pi\phi}\frac{\sin(\pi N\phi)}{\sin(\pi\phi)} = \sqrt{2\pi N F_N(\phi)}\; e^{j(N+1)\pi\phi}$$

故

$$\frac{1}{\sqrt{N}}\sum_{t=1}^{N} x_t e^{-j2\pi f t} = \int_{-1/2}^{1/2} \sqrt{2\pi}\; F_N^{1/2}(\theta-f)\, e^{j(N+1)\pi(\theta-f)}\, dZ(\theta)$$

该式模的平方是式(4-1-14)所定义的周期图，即

$$\tilde{I}_N(f) = \frac{2}{N}\left|\sum_{t=1}^{N} x_t e^{-j2\pi f t}\right|^2$$

$$= 2\int_{-1/2}^{1/2}\int_{-1/2}^{1/2} 2\pi F_N^{1/2}(\theta-f)\, e^{j(N+1)\pi(\theta-f)} \cdot F_N^{1/2}(\theta'-f)\, e^{-j(N+1)\pi(\theta'-f)}\, dZ(\theta)\, dZ^*(\theta')$$

由于$Z(\theta)$是正交过程，$\theta\neq\theta'$时$E[dZ(\theta)dZ(\theta')]=0$，
$E[|dZ(\theta)|^2]=G(\theta)d\theta$（$G(\theta)$是$x_t$的功率谱密度）。故上式的期望为

$$E[\tilde{I}_N(f)] = 2\int_{-1/2}^{1/2} 2\pi F_N(\theta-f)\, E[|dZ(\theta)|^2]$$

$$= 2\int_{-1/2}^{1/2} 2\pi F_N(\theta-f)\, G(\theta)\, d\theta$$

由于$N\to\infty$时$2\pi F_N(\theta)\longrightarrow\delta(\theta)$，故上式右边收敛于$2G(f)$。因

此采用和式(4-1-14)略有不同的定义式：

$$I_N(f)=\frac{1}{2}\tilde{I}_N(f)=\frac{1}{N}\left|\sum_{t=1}^{N}x_t e^{-j2\pi ft}\right|^2 \qquad (4\text{-}2\text{-}1)$$

这时$I_N(f)$将是$G(f)$的渐近无偏估计，即

$$E[I_N(f)]\xrightarrow[N\to\infty]{}G(f)$$

称$I_N(f)$为“修正周期图”以区别于$\tilde{I}_N(f)$，但在不致引起误解时$I_N(f)$也简称周期图。给定N个观察值$x_1\cdots x_N$，可以由式(4-2-1)估计 $|f|\leqslant\frac{1}{2}$ 内的非归一化功率谱 $G(f)$，若再除以样本方差就能得出归一化功率谱$S(f)$的估计。

除了直接从观察数据定义周期图 $I_N(f)$ 外，也可以从过程的自协方差函数来定义$I_N(f)$。如在实过程的情况下，由于

$$G(f)=\sum_{r=-\infty}^{\infty}R(r)\cos 2\pi rf \qquad -\frac{1}{2}\leqslant f\leqslant\frac{1}{2}$$

其中$R(r)=E[x_t x_{t+r}]$，($E[x_t]=0$) 自然地定义

$$I_N(f)=\sum_{r=-(N-1)}^{N-1}\hat{R}(r)\cos 2\pi rf \qquad (4\text{-}2\text{-}2)$$

其中 $\hat{R}(r)=\frac{1}{N}\sum_{t=1}^{N-|r|}x_t x_{t+|r|}$，对式(4-2-2)两边取期望，并

利用 $R(r)=\int_{-1/2}^{1/2}G(\theta)\cos 2\pi r\theta d\theta$可导出

$$E[I_N(f)]=\sum_{r=-(N-1)}^{N-1}E[\hat{R}(r)]\cos 2\pi rf$$

$$=\sum_{r=-(N-1)}^{N-1}\left(1-\frac{|r|}{N}\right)R(r)\cos 2\pi rf$$

$$= \int_{-1/2}^{1/2} G(\theta)\left[\frac{1}{2}\sum_{r=-(N-1)}^{N-1}\left(1-\frac{|r|}{N}\right)\right.$$

$$\left.(\cos 2\pi r(\theta+f)+\cos 2\pi r(\theta-f))\right]d\theta$$

$$= \int_{-1/2}^{1/2} G(\theta)[\pi F_N(\theta+f)+\pi F_N(\theta-f)]d\theta$$

$$= \int_{-1/2}^{1/2} 2\pi F_N(\theta-f)G(\theta)d\theta \qquad (4\text{-}2\text{-}3)$$

从上式最后结果看出它正是$\frac{1}{2}E[\tilde{I}_N(f)]$，从而说明式(4-2-2)同式(4-2-1)是等效的然而是间接地从样本自协方差函数作功率谱的估计。

2. 样本周期图的方差 上面虽然给出了两个估计功率谱密度的计算式(4-2-1)和(4-2-2)，并指出当数据量N很大时估计是无偏的，但这只是周期图统计性质的一个方面。自然地，我们还希望它也是一致的估计，即估计方差会随着N的增大而趋于零。在第二章中曾指出，式(4-2 2)中所取的$\hat{R}(r)$在N很大时是$R(r)$无偏的和一致的估计，那么由此得出的$I_N(f)$是否也是$G(f)$的一致估计呢？事实不然，下面的分析将会看出：尽管周期图是功率谱的渐近无偏估计，其估计方差却依然很大。因此在应用周期图时必须附加某种"改造"手段来减小功率谱密度的估计方差。

下面讨论周期图的方差将要得出的结论是：当x_t为零均值正态平稳过程时，

$$\lim_{N\to\infty} \mathrm{Var}[I_N(f)]=G^2(f) \qquad (4\text{-}2\text{-}4)$$

为证明这一关系，先写出在频率f_1和f_2处周期图的协方差。

$$\mathrm{Cov}[I_N(f_1),\ I_N(f_2)]=E[(I_N(f_1)-E[I_N(f_1)])(I_N(f_2)-E[I_N(f_2)])]$$

$$=E[I_N(f_1)I_N(f_2)]-E[I_N(f_1)]E[I_N(f_2)] \qquad (4\text{-}2\text{-}5)$$

由定义式(4-2-1)，并利用式(3-1-14a)可以导出

$$E[I_N(f_1)I_N(f_2)]$$

$$=\frac{1}{N^2}\sum_t\sum_h\sum_u\sum_v E[x_t x_h x_u x_v]e^{-j2\pi f_1(t-h)} e^{-j2\pi f_2(u-v)}$$

$$=E[I_N(f_1)]E[I_N(f_2)] + \left|\frac{1}{N}\sum_t\sum_u R(t-u)e^{-j2\pi(f_1 t+f_2 u)}\right|^2 + \left|\frac{1}{N}\sum_t\sum_v R(t-v)\, e^{-j2\pi(f_1 t-f_2 v)}\right|^2 \quad (4\text{-}2\text{-}6)$$

式中Σ号只标出变量，其求和范围均为1到N。式(4-2-6)最右边第二项内的双重和式可写成$\sum_u\left[\sum_t R(t-u)e^{-j2\pi f_1 t}\right]e^{-j2\pi f_2 u}$如再引入函数$d_t$：

$$d_t=\begin{cases}1 & 1\leqslant t\leqslant N\\ 0 & 其他\end{cases} \qquad (4\text{-}2\text{-}7)$$

则有

$$\sum_t R(t-u)e^{-j2\pi f_1 t}=\sum_{t=-\infty}^{\infty} d_t R(t-u)e^{-j2\pi f_1 t}$$

这是d_t和 $R(t-u)$ 乘积的富氏变换，它等于各自富氏变换的卷积。将d_t的富氏变换记作$D(f)$，于是

$$\sum_u \left[\sum_t R(t-u)\ e^{-j2\pi f_1 t}\right] e^{-j2\pi f_2 u}$$

$$=\sum_u \left[\int_{-1/2}^{1/2} G(\theta)D(f_1-\theta)\ e^{-j2\pi\theta u} d\theta\right] e^{-j2\pi f_2 u}$$

$$=\int_{-1/2}^{1/2} G(\theta)D(f_1-\theta)\left[\sum_u e^{-j2\pi(\theta+f_2)u}\right] d\theta$$

$$=\int_{-1/2}^{1/2} G(\theta)D(f_1-\theta)D(\theta+f_2)d\theta$$

式(4-2-6)右边第三项等于将第二项的f_2反号，故利用这些结果，并将式(4-2-6)代入式(4-2-5)便得

$$\begin{aligned}
&\mathrm{Cov}[I_N(f_1),\ I_N(f_2)] \\
&= \left| \int_{-1/2}^{1/2} G(\theta)\frac{1}{N}D(f_1-\theta)\ D(\theta+f_2)d\theta \right|^2 \\
&\quad + \left| \int_{-1/2}^{1/2} G(\theta)\frac{1}{N}D(f_1-\theta)D(\theta-f_2)d\theta \right|^2
\end{aligned} \tag{4-2-8}$$

令$f_1=f_2=f$并考虑到

$$\begin{aligned}
D(f) &= \sum_{t=-\infty}^{\infty} d_t e^{-j2\pi ft} = \sum_{t=1}^{N} e^{-j2\pi ft} \\
&= e^{-j\pi(N+1)f}\left[\frac{\sin(\pi fN)}{\sin(\pi f)}\right]
\end{aligned} \tag{4-2-9}$$

$$D(f-\theta)D(\theta-f)=\frac{\sin^2[\pi(f-\theta)N]}{\sin^2[\pi(f-\theta)]}=2\pi NF_N(f-\theta)$$

可得

$$\text{Var}[I_N(f)] = \left| \int_{-1/2}^{1/2} G(\theta)\frac{1}{N} D(f-\theta)D(f+\theta)d\theta \right|^2$$

$$+ \left| \int_{-1/2}^{1/2} G(\theta)2\pi F_N(f-\theta)d\theta \right|^2$$

$$= \left| \int_{-1/2}^{1/2} G(\theta)\frac{1}{N} D(f-\theta)D(f+\theta)d\theta \right|^2$$

$$+ \{E[I_N(f)]\}^2 \tag{4-2-10}$$

由此可见，不论N多大，周期图的方差总等于或大于其期望值的平方，当$N\to\infty$时有关系式(4-2-4)。

式(4-2-9)中两个正弦的比$\sin(\pi fN)/\sin(\pi f)$常称为狄黑赫利(*Dirichlet*)核，记作 $D_N(f)$。显然，$D_N{}^2(f)$ 就是$2\pi NF_N(f)$ (F_N为费泽核)。图4-5为$D_N(f)$曲线。它和费泽核一样都对称于原点，且在原点附近有较大的峰值，通常把在函数为零的第一个正频率和第一个负频率之间的部分称为“主瓣”，其余部分为“旁瓣”，$D_N(f)$ 主瓣宽度为$2/N$，旁瓣正负相间，$F_N(f)$ 主瓣宽度相同，但旁瓣总是正的(见图4-1)。

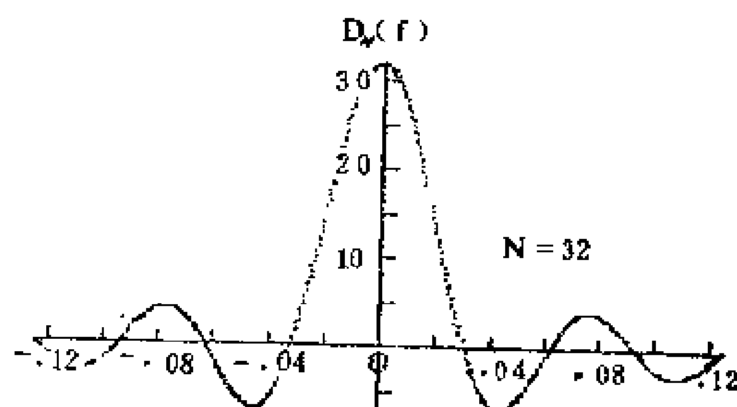


图4-5　狄里赫利核曲线

如果N是足够大的有限值，则$D(f)$的主瓣宽度相当窄（如图

4-6所示，图中将它近似为宽度2B 的矩形)，使得在 $B=a/N<|f|<1/2$范围内$D(f)\simeq 0$。这样，在$B<f<\frac{1}{2}-B$内$D(f-\theta)D(f+\theta)\simeq 0$，那末

$$\mathrm{Var}[I_N(f)]\simeq\{E[I_N(f)]\}^2 \qquad B<f<\frac{1}{2}-B \quad (4-2-11)$$

如果f_1，f_2均在$(0,\ \frac{1}{2}-B)$区间内，则当$|f_2-f_1|>2B$时，$\mathrm{cov}[I_N(f_1),\ I_N(f_2)]\simeq 0$。这表明频率间隔超过 2B 的周期图纵坐标之间实际上是不相关的，由于N很大时B 相应很小，因此频率相

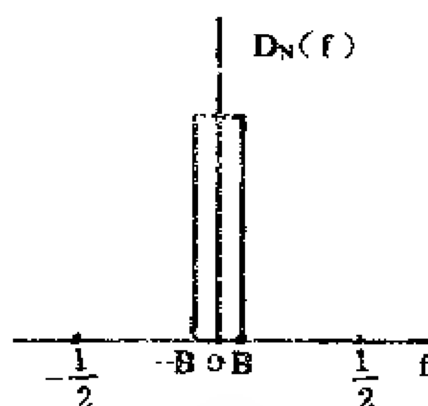


图4-6 窗宽与周期图方差关系说明

距不远的周期图的值就不相关了，这反映在周期图的曲线“抖动”较大。

下面作为一个例子来看正态白噪声序列 x_t 的周期图的方差。设x_t的真正功率谱$G(f)=\sigma x^2$，则由式(4-2-8)有

$$\mathrm{Cov}[I_N(f_1),\ I_N(f_2)]$$

$$=\sigma_x^4\left|\int_{-1/2}^{1/2}\frac{1}{N}\ D(f_1-\theta)D(\theta+f_2)d\theta\right|^2$$

$$+\sigma_x^4\left|\int_{-1/2}^{1/2}\frac{1}{N}D(f_1-\theta)D(\theta-f_2)d\theta\right|^2 \quad (4\text{-}2\text{-}12)$$

利用公式[9]：

$$\int_{-1/2}^{1/2}A(\theta)B^*(\theta)d\theta=\sum_{t=-\infty}^{\infty}a_t b_t^*$$

（式中$A(\theta)$，$B(\theta)$分别为a_t和b_t的富氏变换），可将式(4-2-12)右边第一项中的积分写成

$$\int_{-1/2}^{1/2}\frac{1}{N}D^*(\theta-f_1)D(\theta+f_2)\,d\theta=\frac{1}{N}\sum_{t=-\infty}^{\infty}d_t^2e^{-j2\pi(f_1+f_2)t}$$

$$=\frac{1}{N}\sum_{t=1}^{N}e^{-j2\pi(f_1+f_2)t}$$

$$=\frac{\sin[\pi(f_1+f_2)N]}{N\sin[\pi(f_1+f_2)]}\quad e^{-j2\pi(f_1+f_2)(N+1)}$$

对式(4-2-12)右边第二项的积分也作类似推导，可得白噪声周期图的协方差为

$$\mathrm{Cov}[I_N(f_1),\ I_N(f_2)]=\sigma_x^4\left\{\left[\frac{\sin[\pi(f_1+f_2)N]}{N\sin[\pi(f_1+f_2)]}\right]^2+\left[\frac{\sin[\pi(f_1-f_2)N]}{N\sin[\pi(f_1-f_2)]}\right]^2\right\}$$

$$=\frac{\sigma_x^4}{N}[2\pi F_N(f_1+f_2)+2\pi F_N(f_1-f_2)] \quad (4\text{-}2\text{-}13)$$

由$f_1=f_2=f$得

$$\mathrm{Var}[I_N(f)]=\sigma_x^4\left[1+\frac{\sin^2(2\pi f N)}{N^2\sin^2(2\pi f)}\right]$$

$$=\sigma_x^4\left[1+\frac{2\pi}{N}F_N(2f)\right] \qquad (4\text{-}2\text{-}14)$$

4.3 功率谱估计的两种基本方法

前面讨论指出，按式(4-2-1)估计功率谱密度是渐近无偏的，然而不是一致的。我们知道，$I_N(f)$和样本自协方差函数$\hat{R}(r)$之间的关系，同$G(f)$和理论自协方差函数$R(r)$之间的关系是一样的，对于每个r，$\hat{R}(r)$是$R(r)$的一致估计。究竟是为什么$\hat{R}(r)$的线性组合—$I_N(f)$却不是$R(r)$的线性组合—$G(f)$的一致估计呢？有人认为$I_N(f)$包含了从$r=0$到$r=N-1$的所有样本自协方差，不论N多大，$I_N(f)$总包含样本自协方差函数的"尾部"—它是由少数几对观察数据求得的，因此这一部分是理论自协方差的很差的估计。当我们用的是自协方差的无偏估计式时，其方差为$O[1/(N-|r|)]$，那末随着r的增大估计方差增大的说法是对的，然而$I_N(f)$用的是$R(r)$的有偏估计，不论r多大，自协方差的估计方差总是$O(1/N)$，因此无法用上述理由解释为什么$I_N(f)$的方差不随N增大而趋于零。应当说，由于$I_N(f)$包含了N个样本自协方差，尽管每个方差均为$O(1/N)$，但N项累加的效果所产生的方差是$O(1)$，正如一个随机变量U，它是N个方差为$1/N$的不相关的$\{U_i\}$之和（即$U=\sum_{i=1}^{N}U_i$），因此$\mathrm{Var}[U]=1$。当然，一般说$\{\hat{R}(r)\}$不是不相关的，但效果基本相同，所以粗浅地说，$\mathrm{Var}[I_N(f)]$之所以不趋于零是因为它包含了"太多"的样本自协方差。从这一概念出发，减少方差的一种方法就是将$\hat{R}(r)$的项数减少，特别是省去和样本自协方差尾部所对应的那些项，由于通

常当 $|r|\to\infty$ 时 $R(r)\to 0$，所以这样做并不会导致偏度误差的增加。

减少项数的最简单做法就是截取，即将式（4-2-2）的求和限改为 $M(<N-1)$ 作为 $G(f)$ 的估计．

$$\hat{G}(f)=\sum_{r=-M}^{M}\hat{R}(r)\cos 2\pi rf \qquad M<N-1 \tag{4-3-1}$$

其中 $\hat{R}(r)$ 仍为

$$\hat{R}(r)=\frac{1}{N}\sum_{t=1}^{N-|r|}x_t x_{t+|r|} \tag{4-3-2}$$

由于 $\hat{G}(f)$ 中只包含 $M+1$ 个样本自协方差，所以可望 $\hat{G}(f)$ 的方差大约是 $\frac{M}{N}\mathrm{Var}[I_N(f)]=O(M/N)$，至于 $\hat{G}(f)$ 的期望

$$E[\hat{G}(f)]=\sum_{r=-M}^{M}\left(1-\frac{|r|}{N}\right)R(r)\cos 2\pi rf \tag{4-3-3}$$

只要随着 $N\to\infty$，$\hat{G}(f)$ 仍将是 $G(f)$ 的渐近无偏估计，如果再加上 M 比 N 小得多的条件，即 $M/N\to 0$，则 $\mathrm{Var}[\hat{G}(f)]\xrightarrow[N\to\infty]{}0$，因此可以得到 $G(f)$ 的一致估计。

式(4-3-1)可以看作是 $\hat{R}(r)$ 乘以矩形函数

$$q(r)=\begin{cases}1 & |r|\leqslant M\\ 0 & |r|>M\end{cases} \tag{4-3-4}$$

后进行富氏变换，图4-7表示了式(4-3-4)的图形，常称之为“矩形窗”，窗口函数还可以取其他形式，由于 $R(r)$ 是偶函数，所以 $q(r)$ 也应当是偶函数，并在 $r=0$ 处为最大，这也是对较为可靠的 $\hat{R}(r)$ 值给予较大的加权，例如三角形的窗函数（如图4-8所示）：

$$q(r)=\begin{cases}1-|r|/M & |r|<M\\ 0 & |r|\geq M\end{cases} \tag{4-3-5}$$

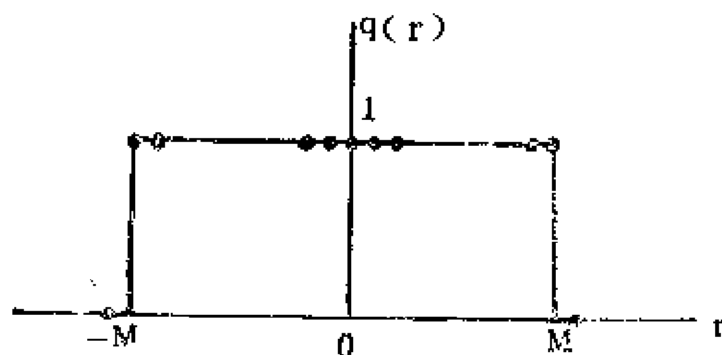


图4-7 矩形窗

这里窗函数是作用在自协方差或自相关函数上，故称所指的窗为“迟后窗”。

一般情况下，带有迟后窗的功率谱估计式为

$$\hat{G}(f)=\sum_{r=-M}^{M}q(r)\hat{R}(r)e^{-j2\pi fr} \tag{4-3-6}$$

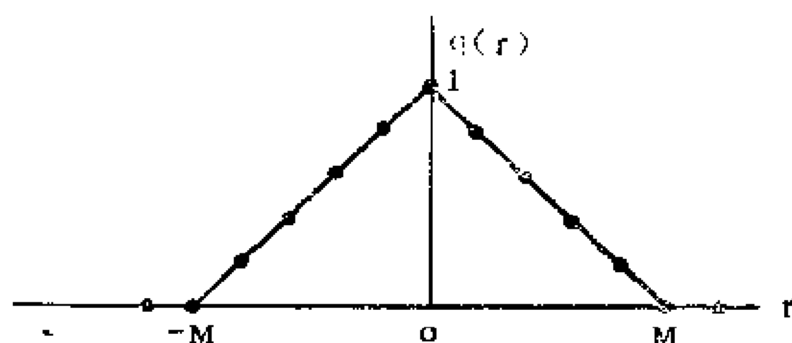


图4-8 三角窗

这种先估计$R(r)$，再对它加窗处理，然后取富氏变换作为功率谱估计的方法称为间接方法。

利用卷积关系，式(4-3-6)可写成

$$\hat{G}(f)=\int_{-1/2}^{1/2}Q(f)I_N(f-\lambda)d\lambda$$

$$= \int_{-1/2}^{1/2} Q(f-\lambda) I_N(\lambda) d\lambda \tag{4-3-7}$$

其中$Q(f)$是$q(r)$的富氏变换，即

$$Q(f) = \sum_{r=-M}^{M} q(r) e^{-j2\pi fr} \tag{4-3-8}$$

作为迟后窗的富氏变换，常称$Q(f)$为“谱窗”。由式（4-3-7)可见，$\hat{G}(f)$是周期图$I_N(f)$通过一个“脉冲响应”为$Q(f)$的“滤波器”产生 的结果，因此$\hat{G}(f)$是经过平滑的周期图。下面再进一步讨论这方法估计的均值和方差。

间接法估计功率谱的期望为

$$E[\hat{G}(f)] = \int_{-1/2}^{1/2} Q(\lambda) E[I_N(f-\lambda)] d\lambda \tag{4-3-9}$$

因$I_N(f)$对于$G(f)$是渐近无偏的，故当N很大时

$$E[\hat{G}(f)] \simeq \int_{-1/2}^{1/2} Q(\lambda) G(f-\lambda) d\lambda \tag{4-3-10}$$

如果谱窗$Q(f)$足够窄，使得$G(f)$在比$Q(\lambda)$主瓣稍宽的范围内可视为常数，则近似有（见图4-9）

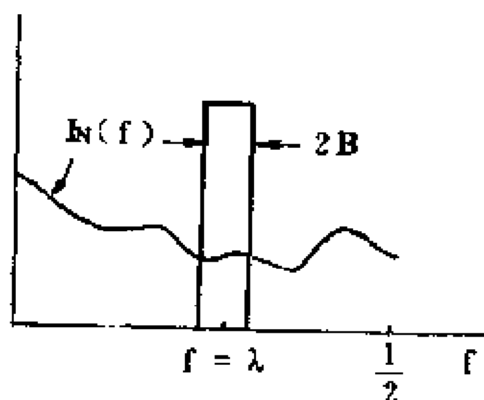


图4-9 谱窗和频谱

$$E[\hat{G}(f)] = G(f) \int_{-1/2}^{1/2} Q(\lambda) d\lambda = G(f) q(0) \tag{4-3-11}$$

显然，无偏的必要条件是

$$q(o)=\int_{-1/2}^{1/2}Q(f)df=1 \tag{4-3-12}$$

$\hat{G}(f)$的协方差为（利用式(4-3-7)和(4-3-10))：

$$\begin{aligned}
\mathrm{Cov}[\hat{G}(f_1),\ \hat{G}(f_2)]&=E[(\hat{G}(f_1)-E[\hat{G}(f_1)])(\hat{G}(f_2)\\
&\quad -E[\hat{G}(f_2)])]\\
&=E[\hat{G}(f_1)\hat{G}(f_2)-\hat{G}(f_1)E[\hat{G}(f_2)]+E[\hat{G}(f_1)]\\
&\quad E[\hat{G}(f_2)]-E[\hat{G}(f_1)]\hat{G}(f_2)]\\
&=E[(\int Q(\lambda)I_N(f_1-\lambda)d\lambda)(\int Q(\mu)I_N(f_2-\mu)d\mu)\\
&\quad -(\int Q(\lambda)I_N(f_1-\lambda)d\lambda)\\
&\quad \cdot(\int Q(\mu)E[I_N(f_2-\mu)]d\mu)+\iint Q(\lambda)Q(\mu)\\
&\quad E[I_N(f_1-\lambda)]E[I_N(f_2-\mu)]d\lambda d\mu\\
&\quad -(\int Q(\lambda)E[I_N(f_1-\lambda)]d\lambda)(\int Q(\mu)I_N(f_2-\mu)d\mu)]\\
&=\iint Q(\lambda)Q(\mu)\{E[I_N(f_1-\lambda)I_N(f_2-\mu)]\\
&\quad -E[I_N(f_1-\lambda)]E[I_N(f_2-\mu)]\}d\mu d\lambda
\end{aligned}$$

上式最后｛·｝内可以证明是$\mathrm{Cov}[I_N(f_1-\lambda),\ I_N(f_2-\mu)]$，故得

$$\mathrm{Cov}[\hat{G}(f_1),\ \hat{G}(f_2)]=\int_{-1/2}^{1/2}\int_{-1/2}^{1/2}Q(\lambda)Q(\mu)\,\mathrm{Cov}[I_N(f_1-\lambda),\ I_N(f_2-\mu)]d\lambda d\mu \tag{4-3-13}$$

为了对式(4-3-13)的意义有具体的了解，设观察数据是来自功率谱$G(f)=\sigma_x^2$的正态白噪声，这时根据式(4-2-13)可以写出

$$\mathrm{Cov}[I_N(f_1-\lambda),\ I_N(f_2-\mu)]=\frac{\sigma_x^4}{N}[2\pi F_N(f_1+f_2-\lambda-\mu)+2\pi F_N(f_1-f_2-\lambda+\mu)]$$

如果N足够大，使得在$-\frac{1}{2}\leqslant\lambda\leqslant\frac{1}{2}$范围内$F_N(f_1+f_2-\lambda-\mu)$和$2\pi F_N(f_1-f_2-\lambda+\mu)$可近似看作为出现在$\lambda=f_1+f_2-\mu$和$\lambda=f_1-f_2+\mu$处的脉冲，则式(4-3-13)为

$$\mathrm{Cov}[\hat{G}(f_1),\ \hat{G}(f_2)]\simeq\frac{\sigma_x^4}{N}\int_{-1/2}^{1/2}Q(\mu)[Q(f_1+f_2-\mu)+Q(f_1-f_2+\mu)]d\mu$$

进一步设$Q(f)$的主瓣足够窄以致可忽略$Q(\mu)Q(f_1+f_2-\mu)$则

$$\mathrm{Cov}[\hat{G}(f_1),\ \hat{G}(f_2)]\simeq\frac{\sigma_x^4}{N}\int_{1/2}^{1/2}Q(\mu)Q(f_1-f_2+\mu)d\mu \tag{4-3-14}$$

对于$|f_1-f_2|$大于主瓣宽度（它近似正比于$1/M$)来说，在f_1和f_2处的两个谱基本值上是不相关的，而若迟后窗宽度较小(M小)，则相邻频率处的谱估值就变得更加相关(即较平滑)。当$f_1=f_2=f$时，由式(4-3-14)得

$$\mathrm{Var}[\hat{G}(f)]\sim\frac{\sigma_x^4}{N}\int_{-1/2}^{1/2}[Q(\mu)]^2d\mu \tag{4-3-15}$$

由于N很大，N个观察值的周期图的方差近似为σ_x^4，所以经过加窗平滑和未经平滑的方差之比为

$$\Lambda=\frac{\mathrm{Var}[\hat{G}(f)]}{\sigma_x^4}=\frac{1}{N}\int_{-1/2}^{1/2}[Q(\mu)]^2d\mu$$

$$=\frac{1}{N}\sum_{r=-M}^{M}q^2(r) \tag{4-3-16}$$

由此亦可看出，当M相对于N减小时，谱估计的方差减小。但考虑到估计的无偏性，M应当大些，因此要作出适度的选择。

应当看到，从加窗后的样本自协方差求富氏变换，有时不能

得到非负的$\hat{G}(f)$，这是间接法估计的主要缺点。

利用相关函数加窗的谱估计程序见附录二中的CMPSD。

功率谱估计的另一种方法是直接法，它是直接对数据序列进行加窗处理，并利用富氏变换来计算 $\hat{G}(f)$ 。事实上，由式(4-2-1）定义的周期图可看作是对x_t加上宽度为N的矩形窗，实际可用的窗形式很多，统称"数据窗"。如果减小谱估计的方差呢？由统计学知，若U_1，…，U_L是无关的随机变量，它们各自具有均值μ和方差σ_u^2，若取这些变量的算术平均，即$\sum_{i=1}^{L} U_i/L$，则均值不变而方差为σ_u^2/L。这一事实启示我们：如果将N点时间序列分成L个不重叠的段落，每段有M个数据，然后求各段的周期图并加以平均则可望使最终得到的谱估计方差减小L倍。这种方法简单说就是对分段数据的周期图作总体平均。此外，对每段数据还可先进行适当的加窗处理。如果数据序列的自相关在迟后小于M时已衰减到可以忽略，则上述各段的周期图可近似认为彼此无关。巴特利用矩形数据窗（即简单分段数据）来完成上述计算，而韦尔契（*Welch*）提出了一些改进，使用了较好的数据窗，并允许各段数据部分重叠。

下面设观察数据为$x(i)$，$i=0$，1，…，$N-1$，将它构成L个M点序列并乘上数据窗$w(i)$，即

$$x_{M,p}(i)=w(i)x(i+pK) \quad p=0,\ 1,\ \cdots,\ L-1;$$
$$i=0,\ 1,\ \cdots,\ M-1 \qquad (4\text{-}3\text{-}17)$$

这里x的下标M表示段长，p表示段号，当$i<0$和$i>M-1$时$w(i)=0$，K为正整数，当$K<M$时各段有部分重叠，$K\geq M$时不重叠。为了使分段后的各数据点不超出N点数据的总长度，应有$(L-1)K+M\leqslant L$。图4-10为分段的示意图。

由式(4-2-1)并考虑到无偏的必要条件（式(4-3-12)），各

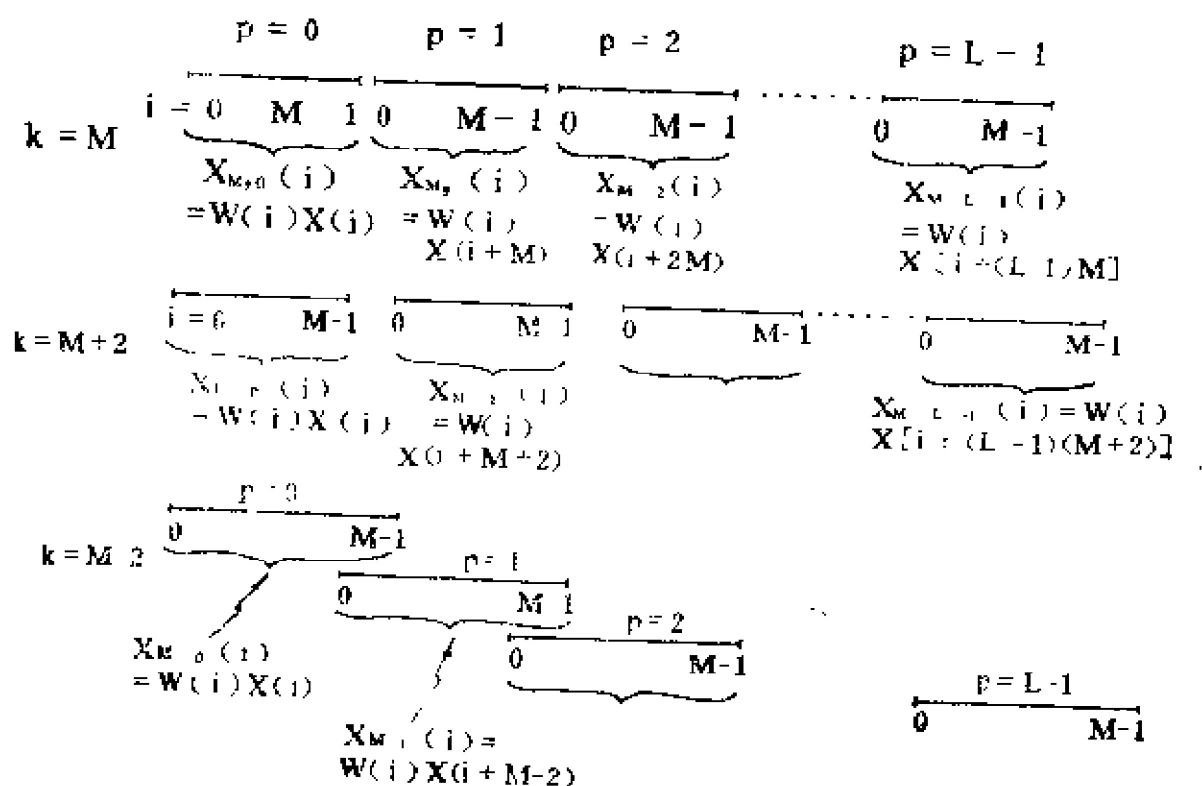


图4-10 数据的分段与重叠

段对应的周期图应为

$$I_{M,p}(f)=\frac{1}{Mq_M(0)}\left|X_{M,p}(f)\right|^2 \quad p=0,1,\cdots L-1 \qquad (4\text{-}3\text{-}18)$$

其中

$$X_{M,p}(f)=\sum_{i=0}^{M-1} w(i)x(i+pK)e^{-j2\pi fi} \qquad (4\text{-}3\text{-}19)$$

$$q_M(0)=\frac{1}{M}\sum_{i=0}^{M-1} w^2(i) \qquad (4\text{-}3\text{-}20)$$

对L个$I_{M,p}(f)$作平均得

$$\hat{G}_M(f)=\frac{1}{L}\sum_{p=0}^{L-1}I_{M,p}(f) \tag{4-3-21}$$

如果 $x(i)$ 是平稳序列，则各段周期图应有相同的均值，$E[\hat{G}_M(f)]$的表达式和式(4-2-3)相同，只是窗函数中的N要换成M，例如用矩形数据窗时有

$$E[\hat{G}_M(f)]=\int_{-1/2}^{1/2}2\pi F_M(f-\theta)G(\theta)d\theta$$

$$=\frac{1}{M}\int_{-1/2}^{1/2}G(\theta)\left[\frac{\sin[\pi(f-\theta)M]}{\sin[\pi(f-\lambda)]}\right]^2d\theta \tag{4-3-22}$$

由于分段后 F_M 的主瓣宽度比未分段时大（例如分段取矩形窗时对应的 F_M主瓣宽为$2/M$，因 $M<N$，故$2/M>2/N$），所以$\hat{G}_M(f)$的偏度误差比不分段时大些。

现在来考虑$\hat{G}_M(f)$的方差。由于

$$\hat{G}_M(f)-E[\hat{G}_M(f)]=\frac{1}{L}\sum_{p=0}^{L-1}\{I_{M,p}(f)-E[I_{M,p}(f)]\} \tag{4-3-23}$$

故

$$\operatorname{Var}[\hat{G}_M(f)]=\frac{1}{L^2}\sum_{p=0}^{L-1}\sum_{q=0}^{L-1}\operatorname{Cov}[I_{M,p}(f),I_{M,q}(f)] \tag{4-3-24}$$

由于$x(i)$是平稳的，上式中的协方差项只取决于$q-p=r$(p，q为段号)，令

$$\Gamma_r(f)\equiv\operatorname{Cov}[I_{M,p}(f),\ I_{M,q}(f)] \tag{4-3-25}$$

将式(4-3-24)改写成

$$L\operatorname{Var}[\hat{G}_M(f)]=\frac{1}{L}\sum_{p=0}^{L-1}\sum_{r=-p}^{-p+(L-1)}\Gamma_r(f)$$

$$=\sum_{\gamma=-(L-1)}^{L-1}\frac{L-|\gamma|}{L}\Gamma_{\gamma}(f)$$

故

$$\mathrm{Var}[\hat{G}_M(f)]=\frac{1}{L}\sum_{\gamma=-(L-1)}^{L-1}\left(1-\frac{|\gamma|}{L}\right)\Gamma_{\gamma}(f)$$

$$=\frac{\Gamma_0(f)}{L}\sum_{\gamma=-(L-1)}^{L-1}\left(1-\frac{|\gamma|}{L}\right)\frac{\Gamma_{\gamma}(f)}{\Gamma_0(f)} \quad (4\text{-}3\text{-}26)$$

其中

$$\Gamma_0(f)=\mathrm{Cov}[I_{M,p}(f),\ I_{M,p}(f)]=\mathrm{Var}[I_M(f)] \quad (4\text{-}3\text{-}27)$$

它是M点序列周期图的方差(未作总体平均)，Γ_{γ}/Γ_0是$I_{M,p}(f)$和$I_{M,q}(f)$ 之间的相关系数。如果不同段落的周期图之间相关性很小，那末L个M 点周期图经算术平均后的方差约为单个M点周期图方差的L分之一，如果M足够大，使得$G_M(f)$在谱窗内变化很小，则对正态过程而言，$\mathrm{Var}[I_M(f)]\cong(E[I_M(f)])^2=(E[I_N(f)])^2\cong\mathrm{Var}[I_N(f)]$，因此当各段周期图不相关时$\hat{G}_M(f)$的方差也大约是$I_N(f)$的$L$分之一，即

$$\mathrm{Var}[\hat{G}_M(f)]=\frac{1}{L}\mathrm{Var}[I_M(f)] \quad (4\text{-}3\text{-}28)$$

当L 很大时方差趋于零，因此这种估计是一致的。由于数据总数N一定的情况下，L 愈大则M愈小，因此一味减小方差会使偏度误差增大，所以也要在得失中权衡。通常先选择足够大的M（如根据谱的分辨力要求，见下节），然后由允许的方差确定应有的N（若各段为互相衔接而不重叠，则$N=LM$）。由于方差和序列的真实谱有关，而后者又属待估的未知函数，往往以零均值正态白噪声的情况作参考。可以证明（可见[10]），对于零均值正态白噪声序列x_t，若将$|X(f)|^2$写成$[R_eX(f)]^2+[I_mX(f)]^2$，则在任一频率处$X(f)$的实部和需部是两个相互独立的正态随机变量，

其均值为零而方差相同，因此任一频率处的功率谱估计是自由度为$\nu=2L$的χ^2分布

$$2L\hat{G}(f)/G(f)\sim\chi^2_{2L}$$

根据所取的α值(如0.05，0.1，0.2)，可按下式求得$G(f)$的置信区间

$$P\left[\frac{2L\hat{G}(f)}{\chi^2_{2L,\alpha/2}}<G(f)<\frac{2L\hat{G}(f)}{\chi^2_{2L,1-\alpha/2}}\right]=1-\alpha \qquad (4\text{-}3\text{-}29)$$

表4-2以$G(f)/\hat{G}(f)$的分贝值列出$1-\gamma=90\%$时平均个数L和相应的置信限。如果精度要求$\pm 2db$，则可取$L=16$进行总体平均。

表4-2

平均数L	4	8	16	32	64	128	256
置信上限db	+4.7	+3.0	+2.0	+1.4	+1.0	+0.7	+0.5
置信下限db	−2.9	−2.2	−1.6	−1.2	−0.8	−0.6	−0.4

在使用直接法作谱估计时，如果计算分段周期图时用到的各段数据有重叠，则表现各段周期图之间相关性的$\Gamma_\gamma(f)$不为零。在N和M一定时，重叠可增加L，而这种相关性和分段数L的增大对估计方差所起的效果是相反的。韦尔契($Welch$)建议在N和M一定条件下，使各段之间有50%的重叠对于降低估计方差是合适的选择，在适当的数据窗情况下，只要数据重叠不太多，分段数的增加有较明显的效果。

图4-11为一$AR(2)$过程的样本分段周期图谱估计的结果对比。它是由1024个数据用直接法进行估计的，a)图是不分段的（即无平均处理）频谱，b)图和c）图分别取$M=256$和$M=32$按韦尔契方法（50%重叠）得出的分段平均结果。上述计算调用了直接法谱估计程序PMPSE(附录二)。

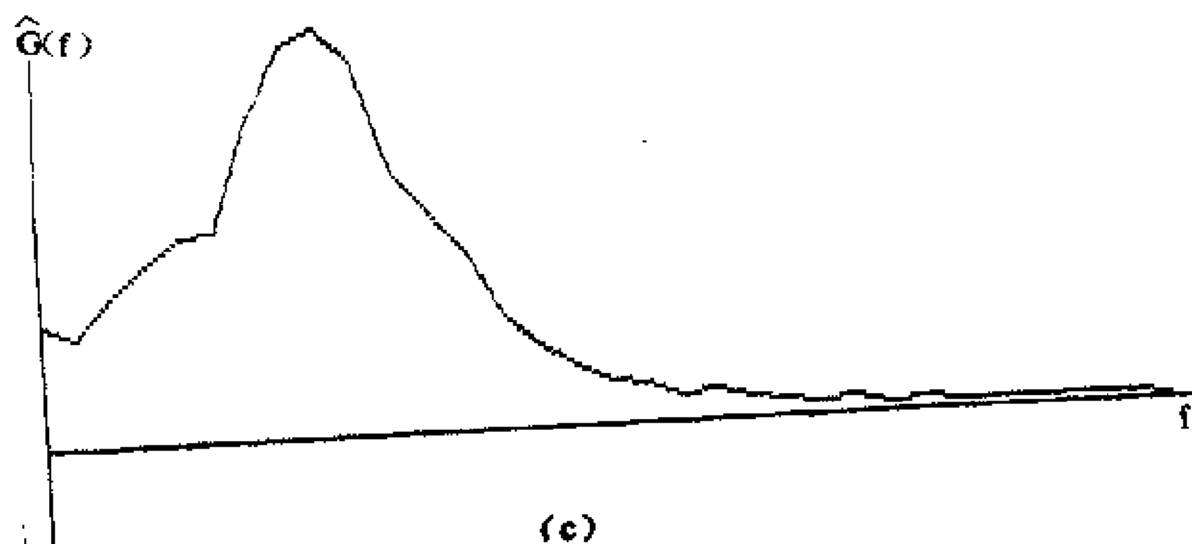
除了对分段周期图进行总体平均外，有时为了进一步减小估计方差，还对总体平均得到的谱估计再和某种形式的谱窗进行卷



(a)



(b)



(c)

图4-11分段周期图的总体平均谱估计

以达到更进一步的平滑。

4.4 窗函数

.1 **窗函数与谱的分辨力和泄漏** 在前面讨论中可看以出，不论从时间序列的数据出发，或是从序列的样本自协方差函数出发估计功率谱，窗函数的引入都是不可避免的，因为最简单的截取就意味着通过了矩形窗。理论上，考虑加窗后谱的估计偏差为

$$E[I_N(f)]-G(f)=G(f)*Q(f)-G(f) \qquad (4\text{-}4\text{-}1)a$$

式中卷积就是式(4-3-10)。如果考虑到迟后窗$q(r)$和数据窗$w(t)$之间有如下关系：

$$q(r)=\frac{1}{N}\sum_{t=-\infty}^{\infty}w(t)w(t+|r|)$$

从而

$$\begin{aligned}Q(f)&=\sum_{r=-\infty}^{\infty}q(r)e^{-j2\pi fr}\\&=\frac{1}{N}\sum_{r=-\infty}^{\infty}\sum_{t=-\infty}^{\infty}w(t)w(t+|r|)e^{-j2\pi fr}\\&=\frac{1}{N}\sum_{t=-\infty}^{\infty}w(t)e^{j2\pi ft}\sum_{r=-\infty}^{\infty}w(t+|r|)e^{-j2\pi f(t+|r|)}\\&=\frac{1}{N}W^{*}(f)W(f)\\&=\frac{1}{N}|W(f)|^{2}\end{aligned}$$

其中$W(f)$是$w(t)$的富氏变换（即所谓“频率窗”）。因此谱估计的偏差亦可用下式表达：

$$E[I_N(f)]-G(f)=G(f)*|W(f)|^{2}/N-G(f) \qquad (4\text{-}4\text{-}1)b$$

然而，实际上要用式（4-4-1）a或b计算偏度误差是难于做到的，因为这里需要用到的真正谱密度 $G(f)$ 本身正是待估的。通常对选择窗函数有指导作用的是两个具有明显意义的因素—分辨力和泄漏。

下面按矩形数据窗的情形加以说明。先看窗的宽度和分辨力的关系。假设序列 x_t 包含两种频率（f_1和f_2）成份，当序列长度无限时，其富氏变换 $X(f)$ 是在f_1和f_2处的脉冲（见图4-12e），将 x_t 加窗后得到的是有限序列 $w_t x_t$，它的富氏变换是脉冲函数和具有狄里赫利核的$W(f)$的卷积，它显然不同于真正的 $X(f)$，矩形窗愈窄（N小）对应的$W(f)$主瓣则愈宽，从而A，B二主瓣互相重叠愈严重（图4-12d）

在使用矩形数据窗时要使二主瓣不重叠，必须 $f_2-1/N\geq f_1+1/N$（这里取$\Delta=1$，否则应将N改为$N\Delta$，下同），即
$\Delta_f=f_2-f_1\geq 2/N$，或者说数据长度（即窗宽）应不小于$2/\Delta_f$，这样可有良好的分辨力。

如果取 $\Delta_f\geq 1/N$或$N\geq 1/\Delta_f$（当$\Delta\neq 1$时为$\Delta_f\geqslant 1/(N\Delta)$ 或 $N\geqslant 1/(\Delta_f\Delta)=2f_{max}/\Delta_f$），则尚可分辨。

由于Δ_f表征了长度为N的序列所能反映的最小频率成份，或者说是能够区分出的最小频率差别，故称Δ_f为有效分辨力带宽。

由于式（4-3-10）卷积的效果可以看成滤波，有限的窗宽可以将谱估计的“细节”平滑掉，使分辨力降低。而数据窗宽→∞时$Q(f)\to\delta(f)$，没有滤波作用，偏度误差为零，分辨力最高，但谱估计的方差必将较大，可能会使高分辨力变得没有意义。因此对分辨力的要求也不能过份。

由于序列长度有限产生的一个问题是“泄漏”。我们知道，有限长度的序列具有无限的带宽，因此在某一特定频率处卷积的结果必然也会使其他频率的成份不同程度地“泄漏”过来，这种现象的严重程度主要取决于谱窗的旁瓣。

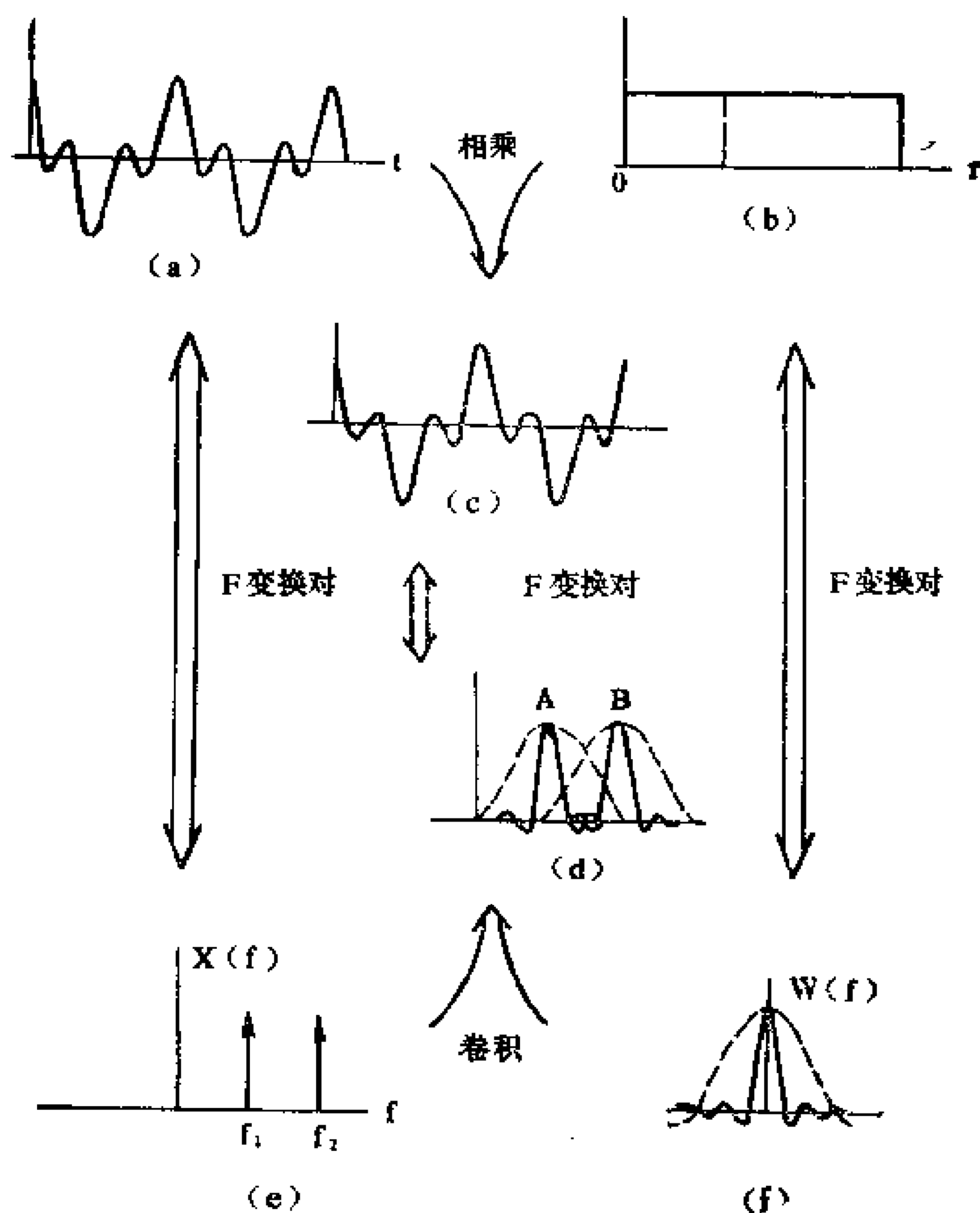


图4-12 窗函数与分辨力

为说明泄漏现象，来考察一个周期为 T_0 的信号 $x(t)$，它包括常值分量和频率为 $1/T_0$ 和 $2/T_0$ 的正余弦分量，因此 $X(f)$ 如图4-13a)所示。如果在信号周期内以 $\Delta=T_0/N$ 均匀间隔采样，而

且信号及其采样点数是无限的，则采样序列的富氏变换是$X(f)$和$p(f)$的卷积，这里$P(f)$是间隔为$1/\Delta$的脉冲

$$P(f)=\frac{1}{\Delta}\sum_{i=-\infty}^{\infty}\delta\left(f-\frac{i}{\Delta}\right) \tag{4-4-2}$$

$X(f)*P(f)$的图像如图4-13c)它是间隔为$1/T_0$的无限脉冲序列。如果以高度为1的矩形窗$w(t)$在一个完整周期范围内（如$0\leqslant t\leqslant T_0$）截取采样序列，这在频率域内等效于$W(f)$和$X(f)*P(f)$的卷积，这里

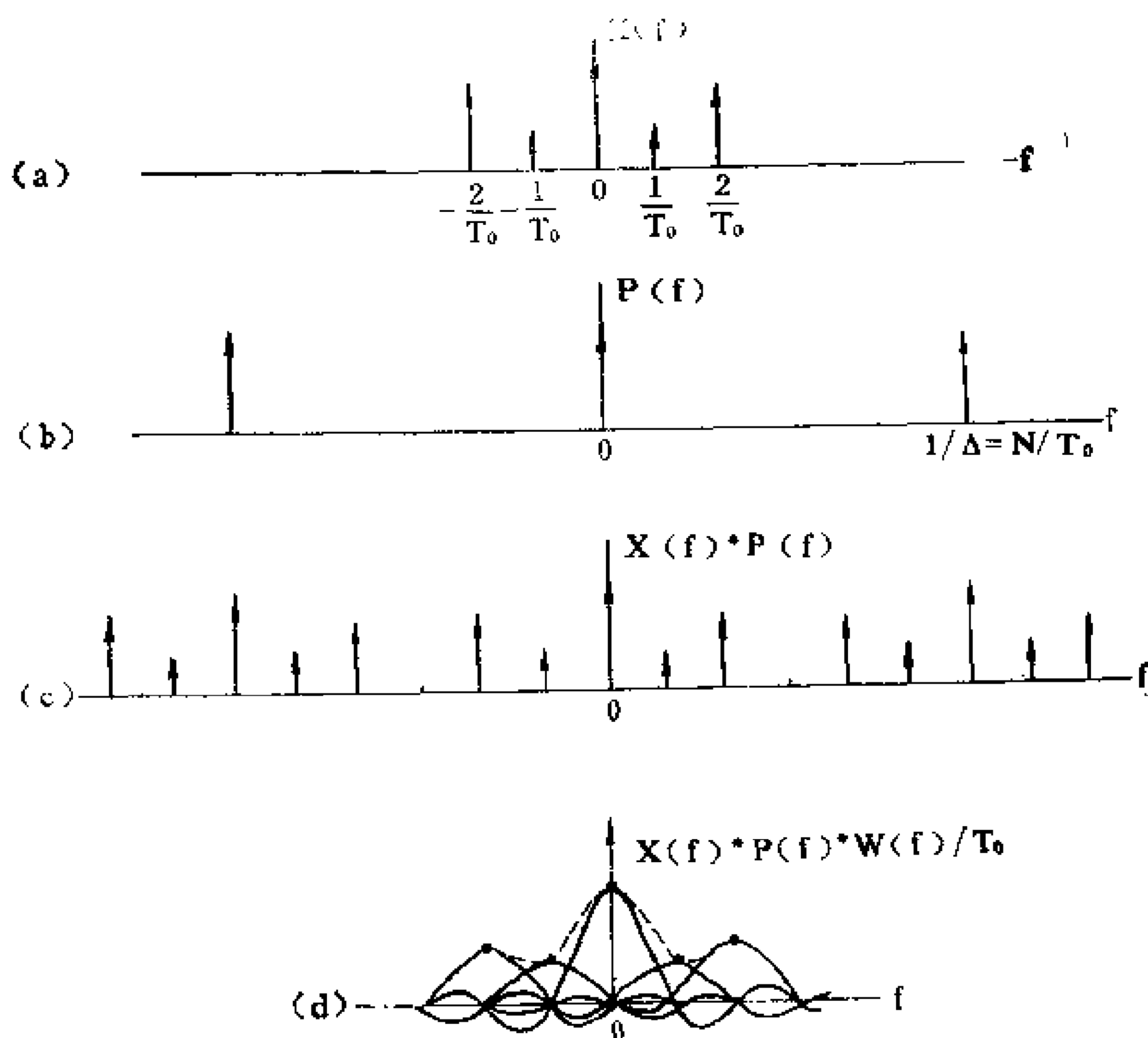


图4-13　泄漏现象图解之一

$$W(f)=F[w(t)]=T_0 e^{-j2\pi f T_0}\frac{\sin(\pi f T_0)}{\pi f T_0} \qquad (4\text{-}4\text{-}3)$$

在$f=\pm k/T_0$处$[X(f)*p(f)*W(f)]/T_0$和原周期信号采样序列的富氏变换相同(注意到$W(f)$在$f=0$处的值为T_0,在其他$\pm k/T_0$处的值为零)，实际上也不难看出，只要窗宽包含序列的若干完整周期，都能在$f=\pm k/T_0$处使采样序列的加窗谱不歪曲其真实谱（除非是混叠效应）.

但是当采样间隔$\Delta=T/N$而$T\neq T_0$时，即便没有混叠效应，由于$W(f)$的旁瓣会在频率为$\pm k/T_0$处使序列加窗后的谱和真正的谱不同（见图4-14），这种现象称为泄漏。要削弱这种泄漏效应,除了主瓣宽度要足够窄外,主要靠减低旁瓣的幅度（正负变号的旁瓣也能部分地抵消泄漏效应）。

上面是以周期信号为例来说明的，实际上往往不可能使采样点在一个周期或它的整数倍的范围内准确地均匀分布，因为周期可能是未知的,而特别是随机过程本身并没有确定的周期,因此泄漏是难免的，必须通过窗函数的适当选择来减少。

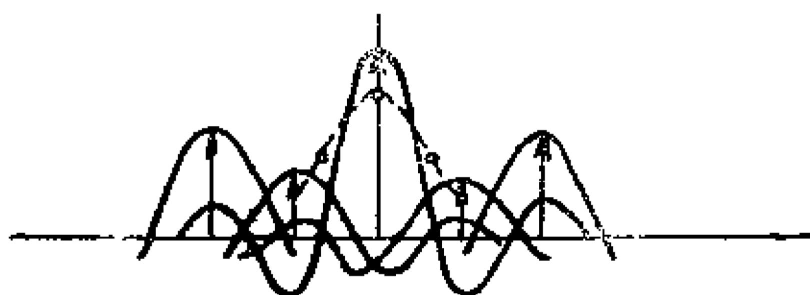


图4-14　泄漏现象图解之二

2．**几种常用的窗函数**　既然窗函数在两种谱估计方法中都是免不了的,而它对谱的分辨力、泄漏及方差都有很大影响,因此有许多人对窗函数的设计及其效果作了研究．本节介绍常用的几

种窗函数，给出其影响谱估计的形状参数。这里给出的窗函数在时间域内以$w(t)$表示，根据使用场合，它既可表示数据窗，亦可表示迟后窗，相应的频率域表达式记作 $W(f)$，它既可为频窗，也可是谱窗。$w(t)$和$W(f)$都取为t和f的连续函数（在窗宽范围内），作数字处理时可改为离散表达式。

窗函数的提出通常基于它的某种最优性质，但有些共同特性可归纳如下：

（1）$w(t)$是实、偶、非负函数。

（2）$W(f)$在原点附近有主瓣，在其两侧有旁瓣。

（3）由于富氏变换具有如下性质：当原函数的m阶导数为脉冲时，象函数的模在$f\to\infty$时趋于f^{-m}。因此若$d^m w(t)/dt^m$出现脉冲，则在半对数坐标图上$|W(f)|$的旁瓣分贝数(db)以 $m\cdot 20$db/dec(dec表示十倍频程）的斜率衰减。

（4）不失一般性，令$w(t)$在原点处为1，且宽度区间是

$-\frac{1}{2}<t<\frac{1}{2}$，即规定

$$w(0)=\int_{-\infty}^{\infty}W(f)df=1 \qquad (4\text{-}4\text{-}4)$$

$$w(t)=0 \qquad |t|\geq\frac{1}{2} \qquad (4\text{-}4\text{-}5)$$

窗函数对估计偏度误差的影响主要通过三个因素：主瓣引起的分辨力下降，邻近旁瓣引起的泄漏和远处旁瓣引起的泄漏。这些因素将用下面四个由$|W(f)/W(0)|$定义的四个形状参数来表征（见图4–15，其中f是按式（4–4–5）条件标准化的）。

a_1—旁瓣的最大峰值

b—主瓣衰减至a_1时的频率

a_2—当旁瓣峰的规律和渐近线吻合时($f=64$)相应的纹波值。

d—旁瓣峰包络线的渐近衰减率

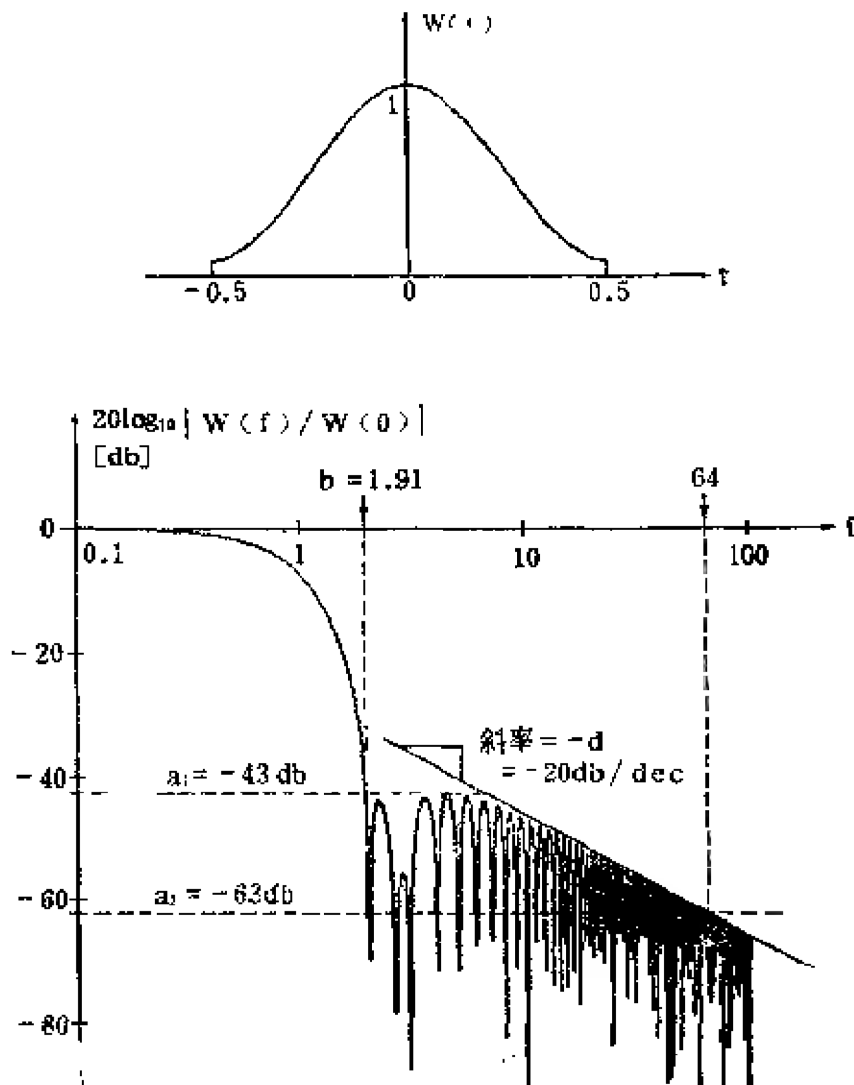


图4-15 以汉明窗为例说明形状参数

图4-15是后面将要提到的汉明(*Hamming*)窗的时域和频域形象及其参数的图示。四个形状参数的意义很明显，*b* 愈小估计的分辨力愈高；a_1愈小则经由邻近旁瓣的泄漏愈少；a_2 愈小和*d*愈大则经由远处旁瓣的泄漏愈少。

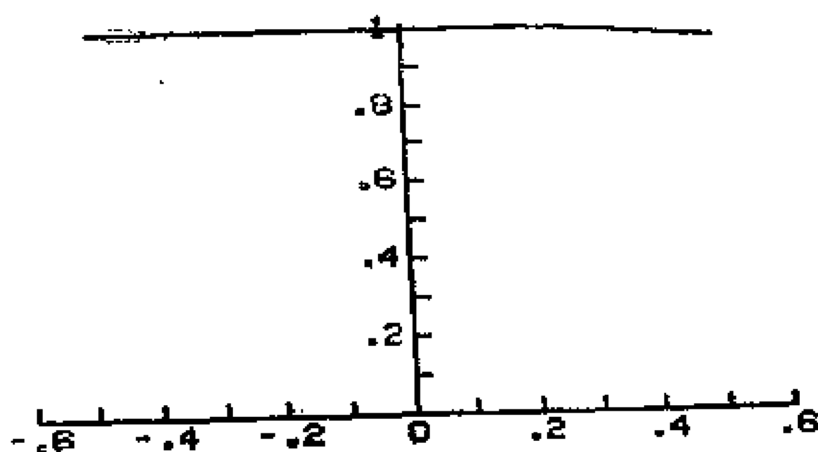
表4-3列出一些窗函数的方程式和参数，在窗名后注有(+)号的表示$W(f)$在所有f处非负。帕波里斯(*Papoulis*)证明了如果谱窗是f的实偶非负函数，且功率谱具有连续二阶导数，则当迟

后窗满足式(4 4-4)(4-4-5)时，估计的偏度误差近似正比于谱窗的二阶矩 $\int_{-\infty}^{\infty} f^2 Q(f)df$，使该二阶矩为最小得出帕波里斯窗，与帕波里斯迟后窗对应的数据窗是余弦边窗。

矩形窗的边缘是突变的，其一阶导数便出现脉冲，故 $d=20$ db/dec。海宁窗用余弦的钟形曲线代替矩形突跳以获得 $60db/dec$ 的渐近衰减率。杜奇建议的窗则是矩形窗和海宁窗的组合。

表4.3

矩形窗



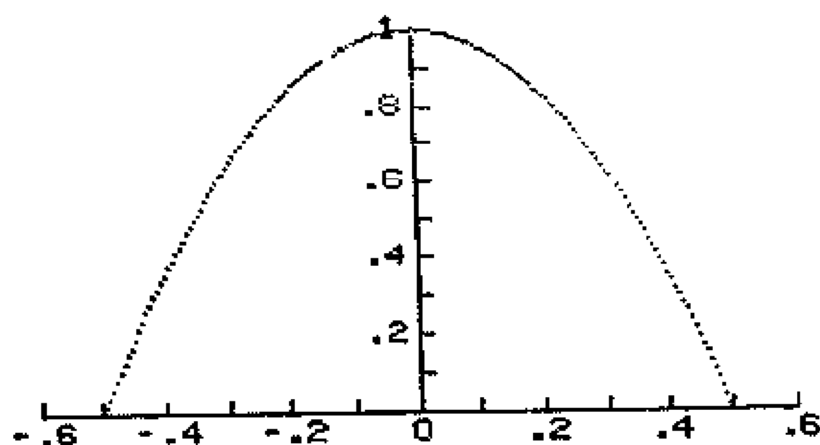
$$w(t)=\begin{cases} 1 & |t|\leqslant 1/2 \\ 0 & \text{其他} \end{cases}$$

$$W(f)=\frac{\sin(\pi f)}{\pi f}$$

$b=0.81$，$a_1=-13\text{db}$，

$a_2=-46\text{db}$，$d=20\text{db/dec}$

帕森-2窗



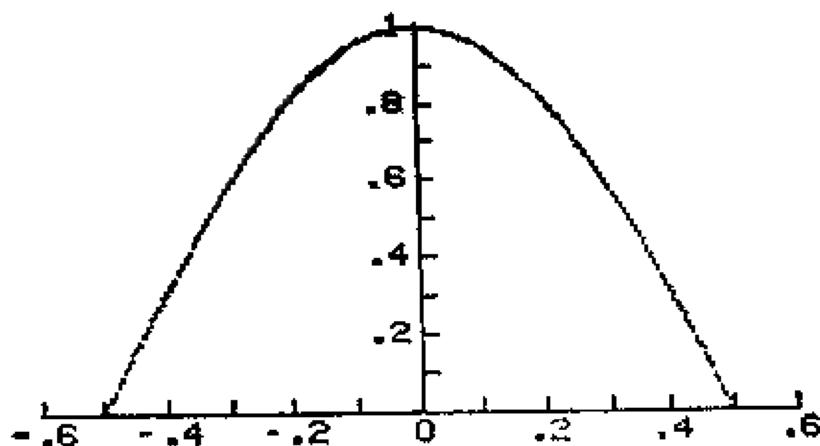
$$w(t)=\begin{cases} 1-4t^2 & |t|\leqslant 1/2 \\ 0 & \text{其他} \end{cases}$$

$$W(f)=\frac{2}{(\pi f)^2}\left[\frac{\sin(\pi f)}{\pi f}-\cos(\pi f)\right]$$

$b=1.28$, $a_1=-21\text{db}$,

$a_2=-83\text{db}$, $d=40\text{db/dec}$

余弦边窗



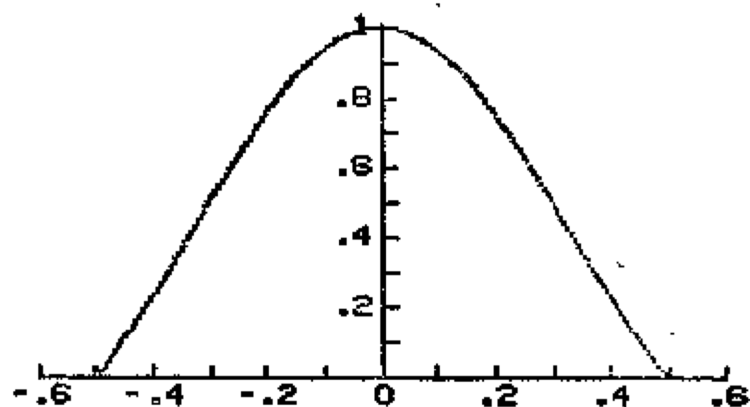
$$w(t)=\begin{cases} \cos(\pi t) & |t|\leqslant 1/2 \\ 0 & \text{其他} \end{cases}$$

$$W(f)=\frac{2\cos(\pi f)}{\pi(1-4f^2)}$$

$b=1.35$, $a_1=-23\text{db}$

$a_2=-84\text{db}$, $d=40\text{db/dec}$

sinx/x窗



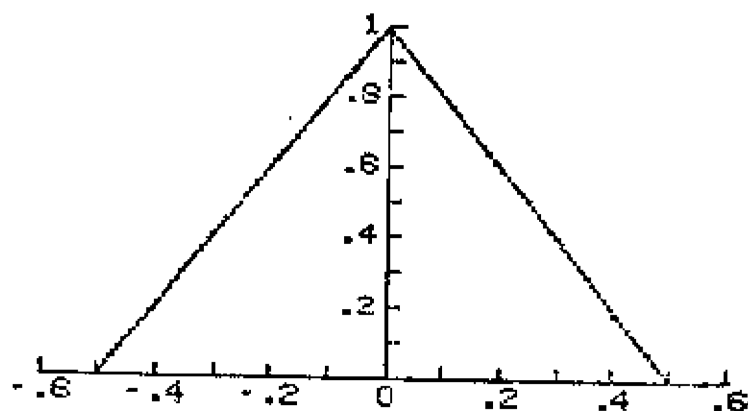
$$w(t)=\begin{cases}\dfrac{\sin(2\pi t)}{2\pi t} & |t|\leqslant 1/2\\ 0 & \text{其他}\end{cases}$$

$$W(f)=\frac{\mathrm{Si}(f+\pi)-\mathrm{Si}(f-\pi)}{2\pi}$$

$b=1.50$, $a_1=-26\text{db}$,

$a_2=-88\text{db}$, $d=40\text{db/dec}$

巴特利(十)窗



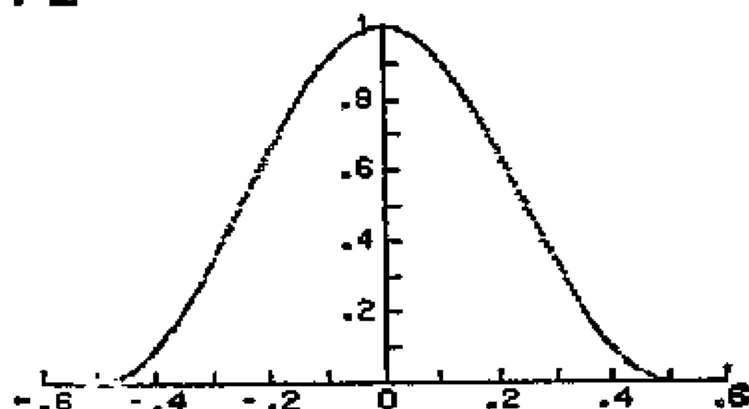
$$w(t)=\begin{cases}1-2|t| & |t|\leqslant 1/2\\ 0 & \text{其他}\end{cases}$$

$$W(f)=\frac{1}{2}\left[\frac{\sin(\pi f/2)}{\pi f/2}\right]^2$$

$b=1.63$，$a_1=-26\text{db}$

$a_2=-80\text{db}$，$d=40\text{db/dec}$

海宁窗



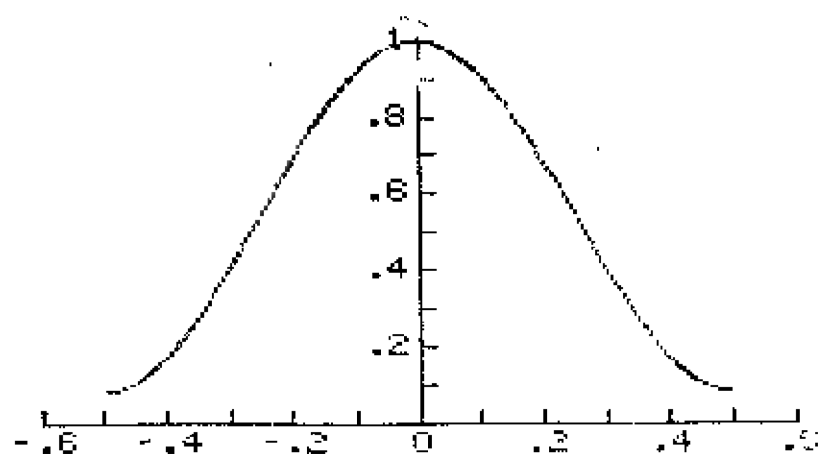
$$w(t)=\begin{cases}\frac{1}{2}+\frac{1}{2}\cos(2\pi t) & |t|\leqslant\frac{1}{2}\\ 0 & \text{其他}\end{cases}$$

$$W(f)=\frac{\sin(\pi f)}{2\pi f(1-f^2)}$$

$b=1.87$，$a_1=-32\text{db}$

$a_2=-118\text{db}$，$d=60\text{db/dec}$

汉明窗



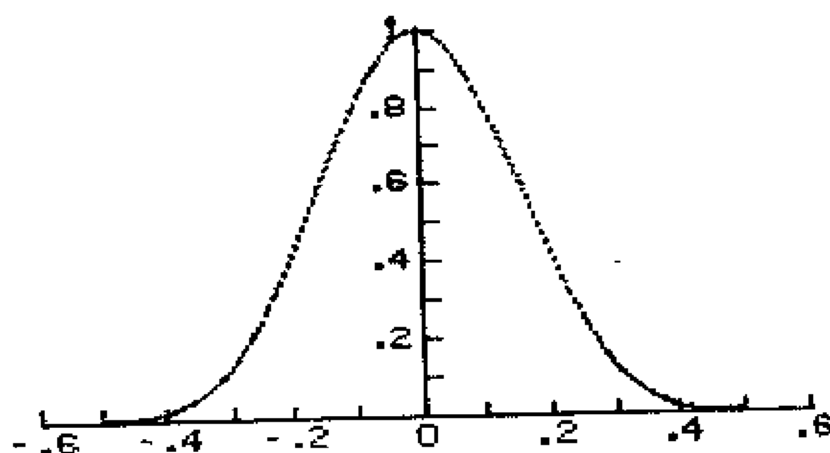
$$w(t)=\begin{cases}0.54+0.46\cos(2\pi t) & |t|\leqslant\frac{1}{2}\\ 0 & \text{其他}\end{cases}$$

$$W(f)=\frac{(1.08-0.16f^2)\sin(\pi f)}{2\pi f(1-f^2)}$$

$b=1.91$，$a_1=-43\text{db}$，

$a_2=-63\text{db}$，$d=20\text{db/dec}$

帕波里斯(十)窗



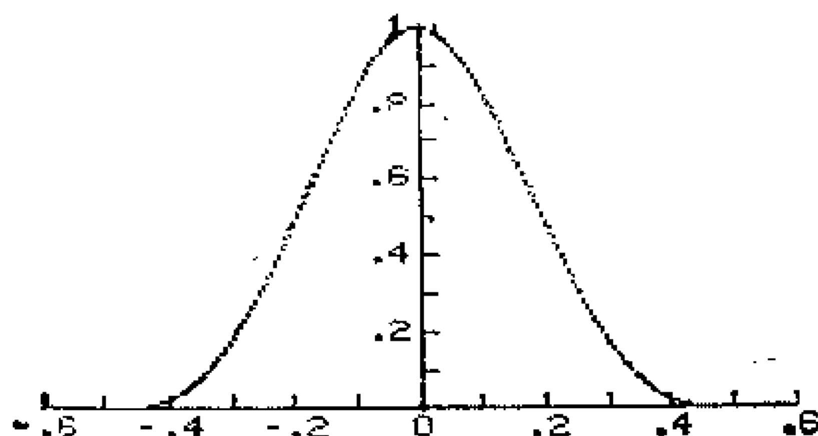
$$w(t)=\begin{cases}\frac{1}{\pi}|\sin(2\pi t)|+(1-2|t|)\cos(2\pi t) \\ \qquad\qquad |t|\leqslant\frac{1}{2} \\ 0 \qquad\qquad 其他\end{cases}$$

$$W(f)=\frac{2+2\cos(\pi f)}{\pi^2(1-f^2)^2}$$

$b=2.70$，$a_1=-46\text{db}$

$a_2=-145\text{db}$，$d=80\text{db/dec}$

伯来克曼窗



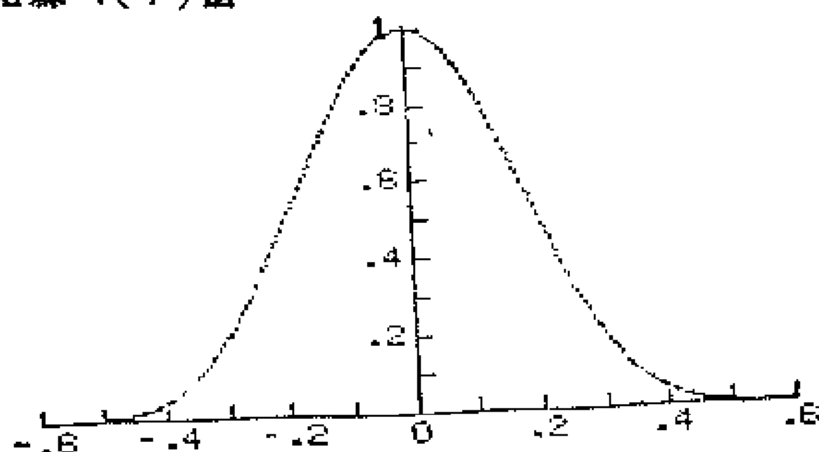
$$w(f)=\begin{cases}0.42+0.50\cos(2\pi t)+0.08\cos(4\pi t) \\ \qquad\qquad |t|\leqslant\frac{1}{2} \\ 0 \qquad\qquad 其他\end{cases}$$

$$W(f)=\frac{(0.18f^2-1.68)\sin(\pi f)}{\pi f(1-f^2)(f^2-4)}$$

$b=2.82 \quad a_1=-58\text{db}$

$a_2=-126\text{db}, \ d=60\text{db/dec}$

帕森-1(+)窗



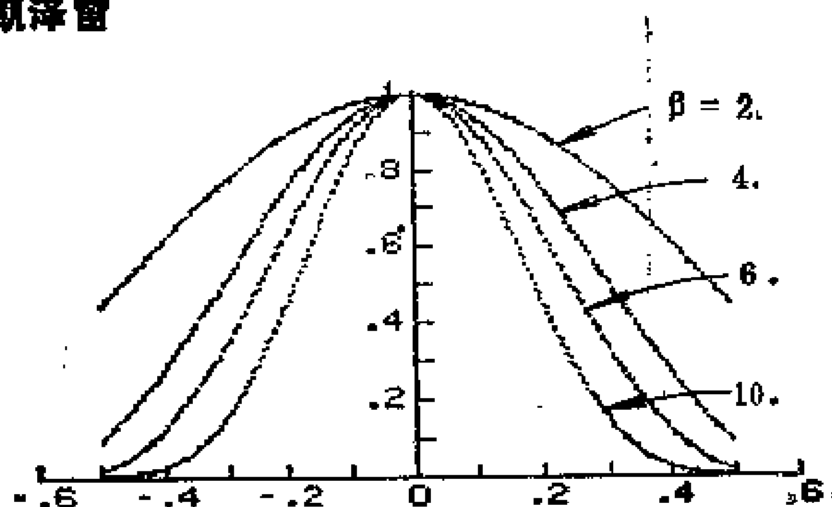
$$
w(t)=\begin{cases} 1-24|t|^2(1-2|t|) & |t|\leqslant\frac{1}{4} \\ 2(1-2|t|)^3 & \frac{1}{4}\leqslant|t|\leqslant\frac{1}{2} \\ 0 & \text{其他} \end{cases}
$$

$$
W(f)=\frac{3}{8}\left[\frac{\sin(\pi f/4)}{\pi f/4}\right]^4
$$

$b=3.25, \ a_1=-53\text{db}$

$a_2=-136\text{db}, \ d=80\text{db/dec}$

凯泽窗

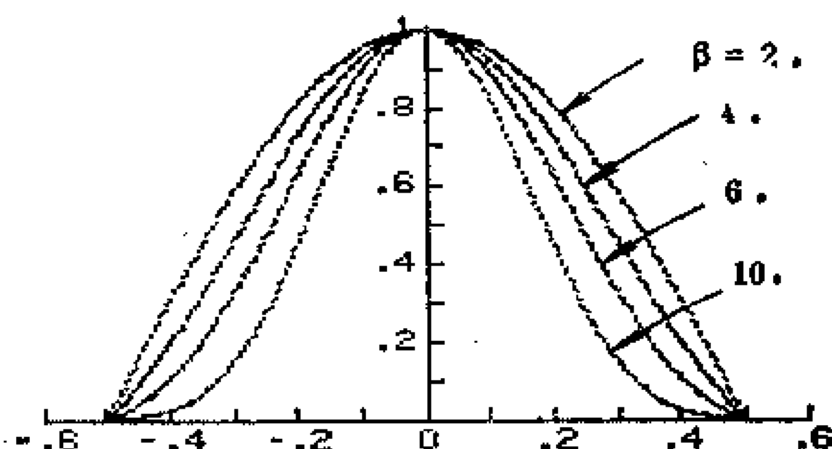


$$w(t)=\begin{cases}\dfrac{I_0(\beta\sqrt{1-(2t)^2})}{I_0(\beta)} & |t|\leqslant\dfrac{1}{2}\\ 0 & 其他\end{cases}$$

$$W(f)=\frac{\sin(\sqrt{\pi^2f^2-\beta^2})}{I_0(\beta)\sqrt{\pi^2f^2-\beta^2}}\qquad 0\leqslant\beta\leqslant 10$$

b，a_1，a_2为β的函数，$d=20$db/dec

修正的凯泽窗



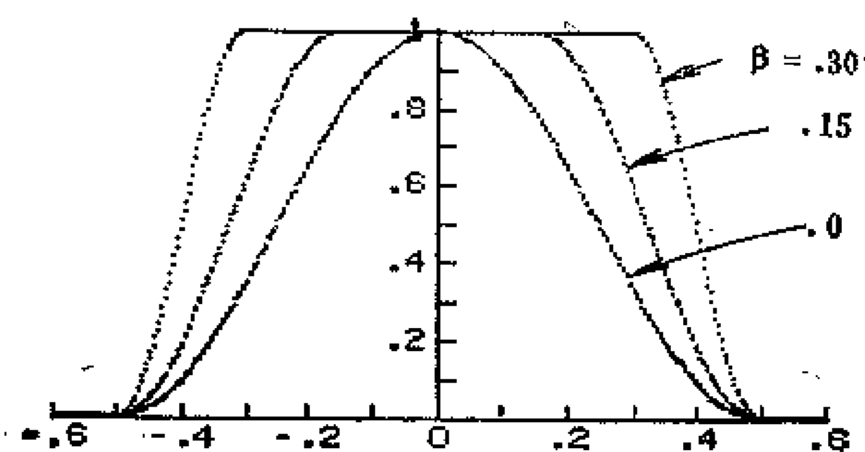
$$w(t)=\begin{cases}\dfrac{I_0(\beta\sqrt{1-4t^2})-1}{I_0(\beta-1)} & |t|\leqslant\dfrac{1}{2}\\ 0 & 其他\end{cases}$$

$$W(f)=\frac{\dfrac{\sin\sqrt{\pi^2f^2-\beta^2}}{\sqrt{\pi^2f^2-\beta^2}}-\dfrac{\sin(\pi f)}{\pi f}}{I_0(\beta)-1}$$

$1 \leqslant \beta \leqslant 10$

b，a_1，a_2为β的函数，$d=40$db/dec

杜奇窗



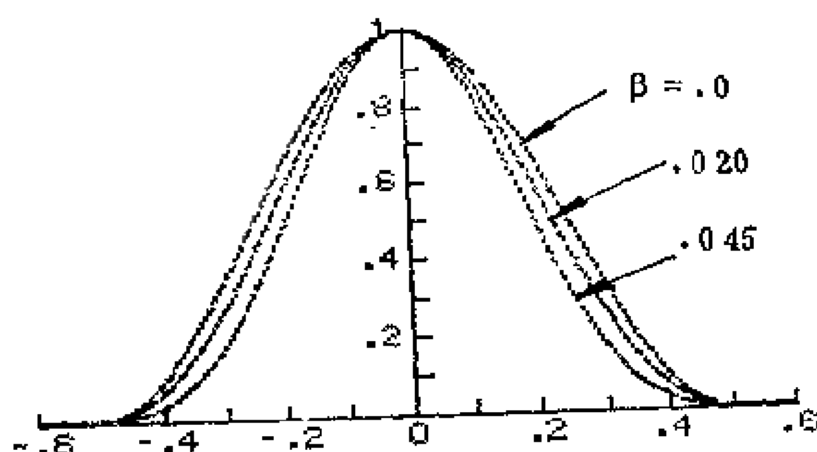
$$w(t)=\begin{cases} 1 & |t| \leqslant \beta \\ \dfrac{1}{2}+\dfrac{1}{2}\cos\left(\dfrac{2\pi(|t|-\beta)}{1-2\beta}\right) & \beta \leqslant t \leqslant \dfrac{1}{2} \\ 0 & \text{其他} \end{cases}$$

$$W(f)=\frac{\sin\left(\dfrac{\pi f(1+2\beta)}{2}\right)\cos\left(\dfrac{\pi f(1-2\beta)}{2}\right)}{\pi f[1-(1-2\beta)^2 f^2]}$$

$0 \leqslant \beta \leqslant 1/2$

b，a_1，a_2为β的函数，$d=60$db/dec

三系数窗



$$w(t)=\begin{cases}\dfrac{1-4\beta}{2}+\dfrac{1}{2}\cos(2\pi t)+2\beta\cos(4\pi t) & |t|\leqslant\dfrac{1}{2}\\ 0 & \text{其他}\end{cases}$$

$$W(f)=\frac{[(1-16\beta)f^2-(4-16\beta)]}{2\pi f(1-f^2)(f^2-4)}\quad\sin(\pi f)$$

$0\leqslant\beta\leqslant0.045$

b，a_1，a_2为β的函数，$d=60$db/dec

注，表中函数$S_i(x)$表正弦积分，即 $S_i(x)=\int_0^x\frac{\sin u}{u}du$，

$I_0(x)$是第一类零阶修正的贝塞尔函数。

时间域内的矩形窗在频率域内的旁瓣衰减缓慢但有正负变化，设想用旁瓣宽度两倍的矩形频窗和它作卷积可使旁瓣平滑（见图4-16），这样做相当于在时间域内将矩形窗乘以 $\sin(2\pi t)/(2\pi t)$，这就引出了 $\sin x/x$ 窗。也可以证明，在 $w(t)$ 的能量一

定，即

$$\int_{-\infty}^{\infty} w^2(t)dt = \int_{-\infty}^{\infty} W^2(f)df = \text{const} \qquad (4\text{-}4\text{-}6)$$

且$w(t)$非负、$W(f)$为实偶函数的条件下，$\sin x/x$窗可使$W(f)$的主瓣在$|f|<1$内的面积为最大。

当数据窗取为矩形时，相应的迟后窗为巴特利窗。从b和a_1看，$\sin x/x$窗比巴特利窗好，但后者的富氏变换是非负的。

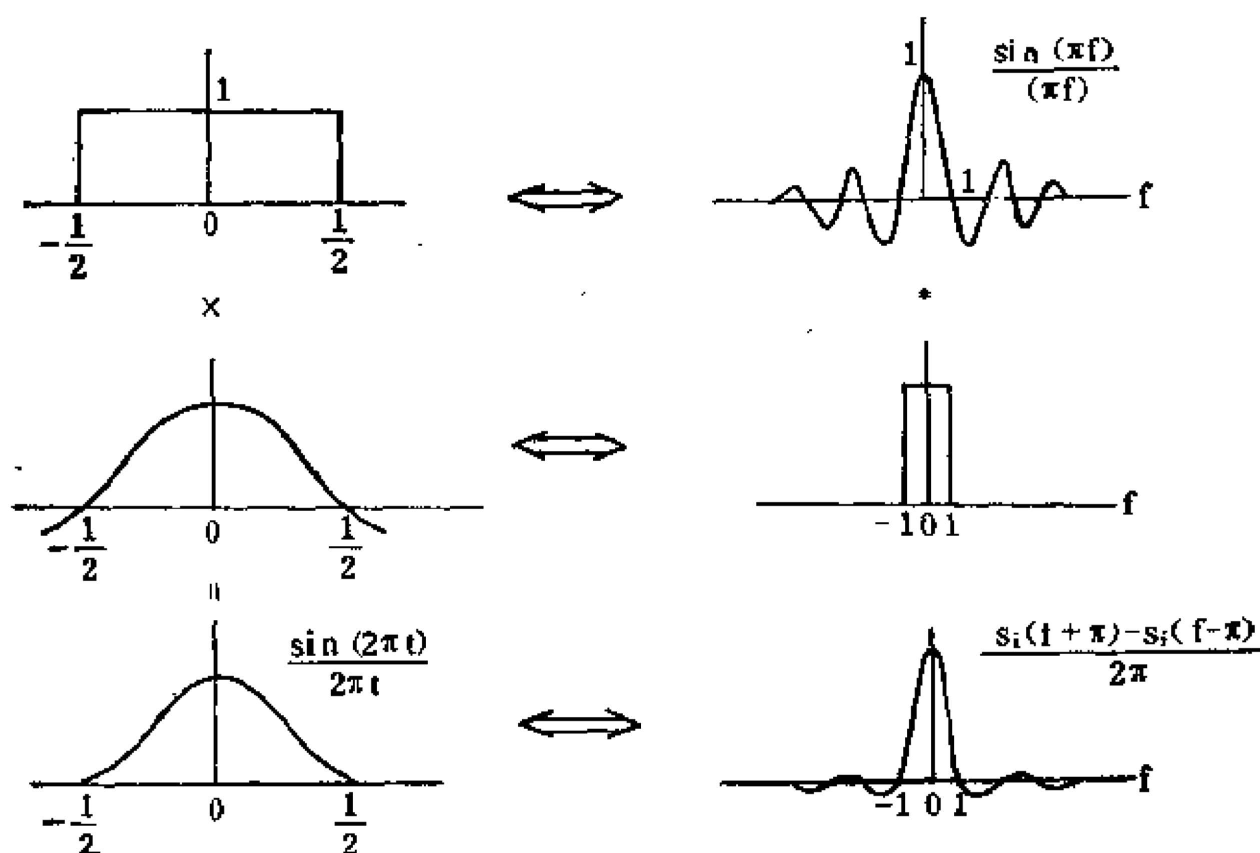


图4-16 $\sin x/x$窗的引出

有一类窗是根据在某一频率区间内的能量对总能量之比为最大来导出，即

$$\int_{-f_1}^{f_1} |W(f)|^2 df \Big/ \int_{-\infty}^{\infty} |W(f)|^2 df = \max \quad (4\text{-}4\text{-}7)$$

凯泽（$Kaiser$）对这一复杂函数作了简化，近似得出凯泽窗函数族，凯泽窗在 $t=\pm 1/2$ 处非零，$d=20db/dec$。为了改善其渐近衰减率，将凯泽窗降低以使其在$t=\pm 1/2$处的值为零，这是利用线性运算：

$$\frac{w(t)-w(1/2)}{1-w(1/2)} \longleftrightarrow \frac{W(f)-\dfrac{w(1/2)\sin(\pi f)}{\pi f}}{1-w(1/2)} \quad (4\text{-}4\text{-}8)$$

由此得到的窗函数族称为修正的凯泽窗。

称为三系数窗的一族窗函数是把$w(t)$在$|t|<1/2$内展成三项富氏级数，它们的衰减率为 60db/dec，当$\beta=0$和0.04 时分别化为海宁窗和伯来克曼($Blackman$)窗。

关于以上各种窗的性质的归纳和类比可参阅有关文献（如参考文献〔11〕），由于各形状参数间的相互依赖关系，要选择一个合适的窗函数并不是直接了当的事情，由实际经验获得的工程直觉可能是无法替代的，当然，研究各种窗的性质不但可以在特定的应用场合缩小选择的范围，而且能够帮助我们权衡利弊。尽管还可以举出其他形式的窗函数，但能 4-3 所列在通常情况下已相当够用了，因此不再列出没有明显改进的窗函数。

上面介绍的窗函数都是用$-\frac{1}{2}\sim\frac{1}{2}$的时间区间内的连续函数表达，在对离散的数据序列或相关函数序列加窗时实际是用窗函数的离散表位式，例如当样点数为N时，窗函数$w(i)$的取值是在$-\frac{N-1}{2}\leqslant i\leqslant\frac{N-1}{2}$（$N$为奇数）或$-\frac{N}{2}\leqslant i\leqslant\frac{N}{2}-1$（$N$为偶数）的范围内。一些离散的窗函数例如

矩形窗 $w(i)=1$ $\quad$ (4-4-9)

巴特利窗

$$w(i)=\begin{cases}1-\dfrac{|2i|}{N+1} & N\text{奇}\\[2ex] 1-\dfrac{|2i+1|}{N} & N\text{偶}\end{cases} \qquad (4\text{-}4\text{-}10)$$

海宁窗

$$w(i)=\begin{cases}0.5+0.5\cos\dfrac{2\pi i}{N+1} & N\text{奇}\\[2ex] 0.5+0.5\cos\dfrac{2\pi(2i+1)}{2(N+1)} & N\text{偶}\end{cases} \qquad (4\text{-}4\text{-}11)$$

汉明窗

$$w(i)=\begin{cases}0.54+0.46\cos\dfrac{2\pi i}{N-1} & N\text{奇}\\[2ex] 0.54+0.46\cos\dfrac{2\pi(2i+1)}{2(N-1)} & N\text{偶}\end{cases} \qquad (4\text{-}4\text{-}12)$$

凯泽窗

$$w(i)=\begin{cases}I_0[\beta\sqrt{1-4i^2/(N-1)^2}]/I_0(\beta) & N\text{奇}\\ I_0[\beta\sqrt{1-4(i+1/2)^2(N-1)^2}]/I_0(\beta) & N\text{偶}\end{cases} \qquad (4\text{-}4\text{-}13)$$

有一种离散窗函数，称道尔夫一切比雪夫(*Dolph-chebyshev*)窗（它没有对应的连续表达式），其特点是：同样N下，在旁瓣不超过a_1的离散窗中具有最窄的主瓣（即参数b最小）。有关这种窗的介绍可参阅有关文献(如参考文献[12])。

在各种时间域的窗函数中，有些是可以用两项或三项复富氏级数表达的，如海宁窗，汉明窗和三系数窗就是这样：

海宁窗—$A_0=0.05$，$A_{-1}=A_1=0.25(A_k=0,\ |k|>1)$

汉明窗—$A_0==0.54$，$A_{-1}=A_1=0.23(A_k=0,\ |k|>1)$

三系数窗—$A_0=0.5-2\beta$，$A_{-1}=A_1=0.25$，$A_{-2}=A_2=\beta$
$(A_k=0,\ |k|>2)$

根据式（A-1-62）关系不难求得这类窗函数的频域表达式。图4-17c 表示以海宁窗作为迟后窗时的情况（取$N=16$为例）及其对应的谱窗为图形（图4-17d）

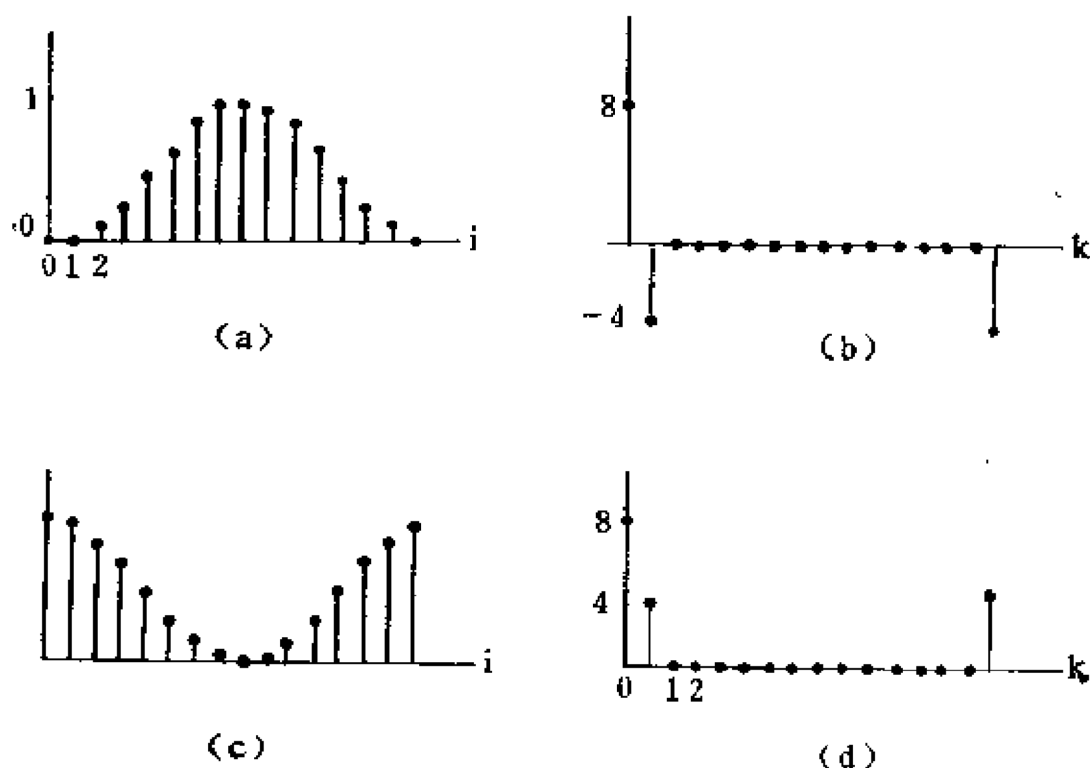
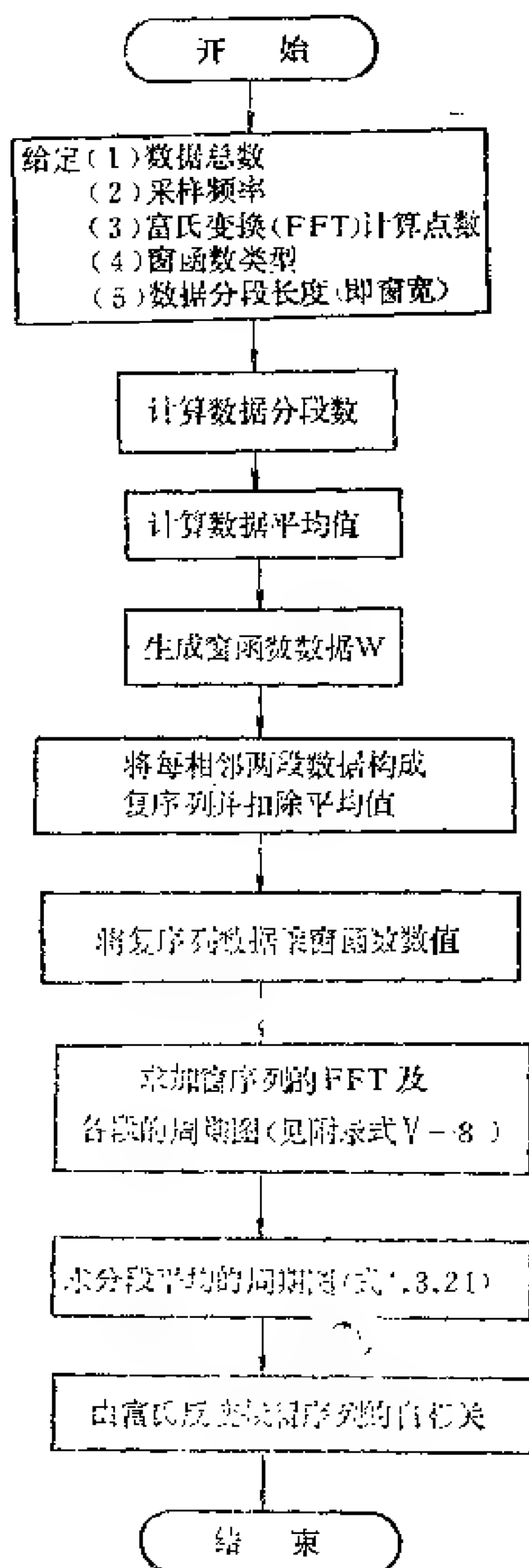


图4-17　连续窗的离散采样及时域窗和频域窗的比较

应当注意，由于数据序列通常是以$\{x_0, x_1, \cdots, x_{N-1}\}$的形式给出，所以数据窗也应在$0\sim N-1$内取值，这时要将前面介绍的窗函数位移 $N/2$。例如以海宁窗作为数据窗时其图形如图4-17a所示。根据位移公式(A-1-29)，移位后的频窗函数应当是谱窗函数乘以$W^{Nk/2}=(-1)^k$，故下标为奇数的频窗是谱窗的反号（图4-17b）。

关于直接法和间接法的谱估计已有一些实用程序可资利用参考文献〔21〕，在实际计算中都要用快速算法求取相关数或数据的富氏变换，有关这些算法的基本原理可参阅附录一的有关内容。



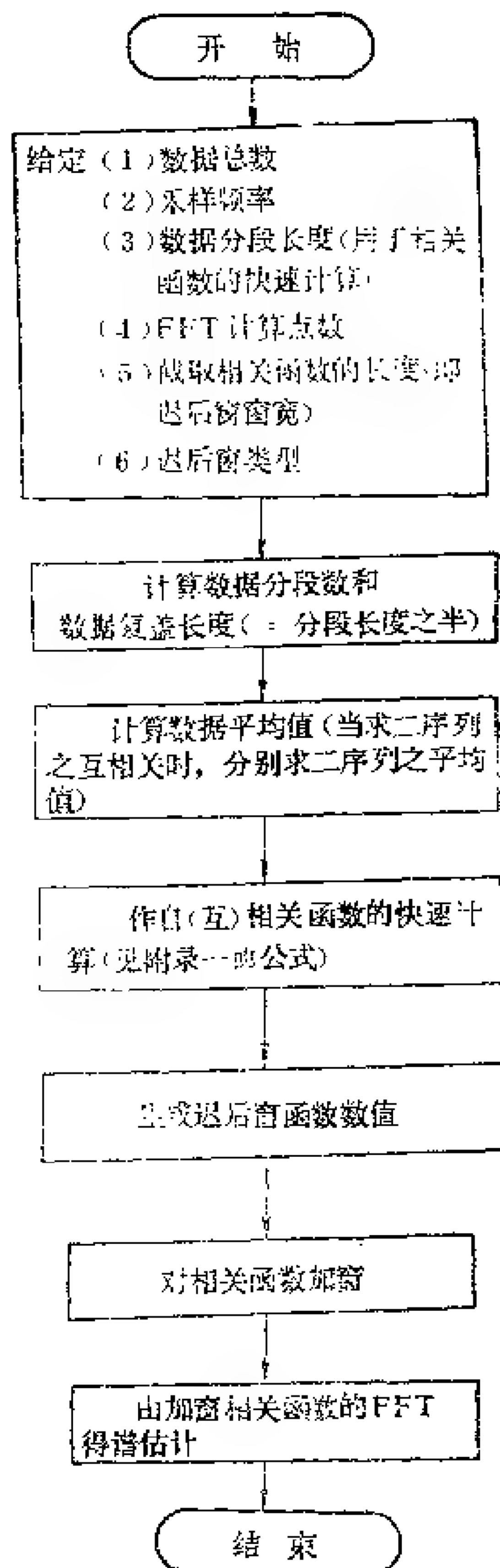


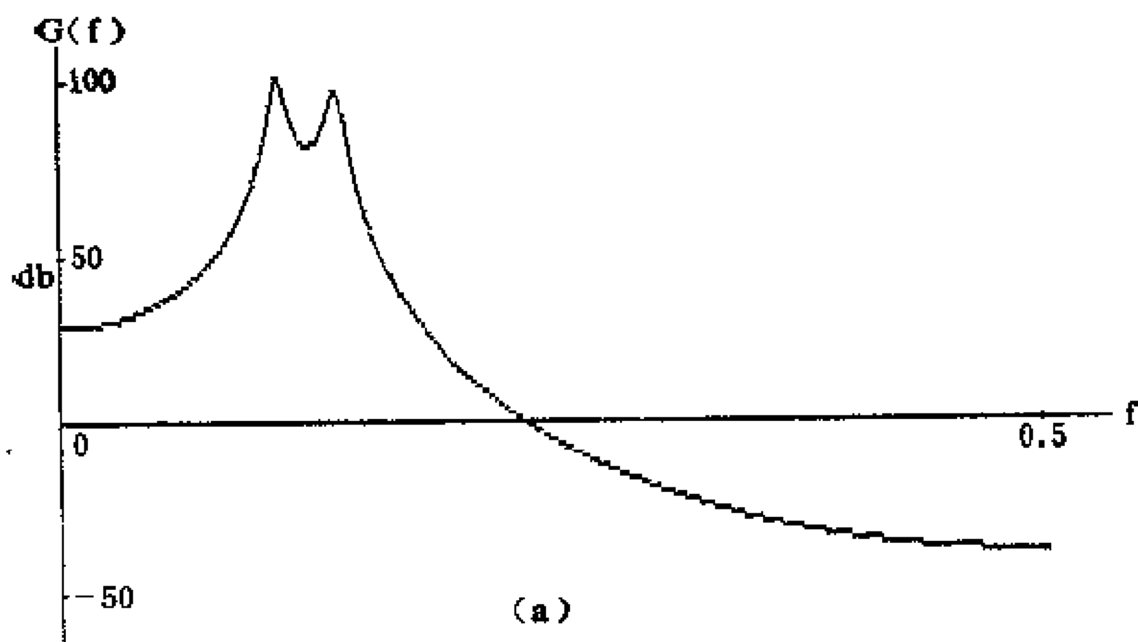
图4-18 直接法谱估计的流程图　　图4-19 间接法谱估计的流程图

图4-18和图 4-19 分别为直接法和间接法谱估计的计算流程图。

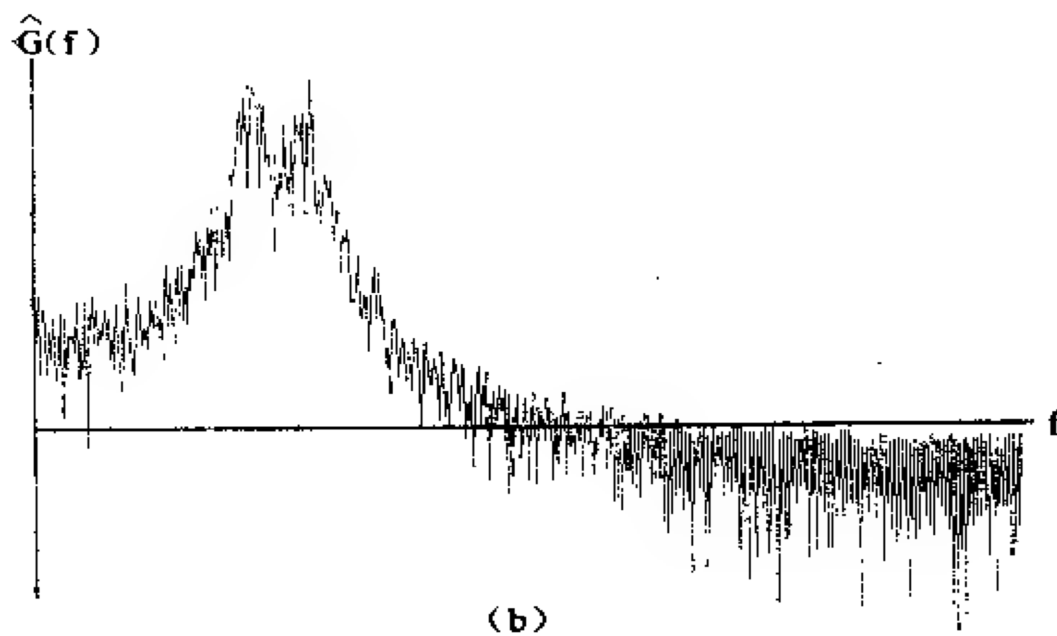
图4-20是一个谱估计计算的例子，它是零均值单位方差白噪声通过传递函数为

$$G(B)=\frac{1}{1-2.7607B+3.8106B^2-2.2635B^3+0.9238B^4}$$

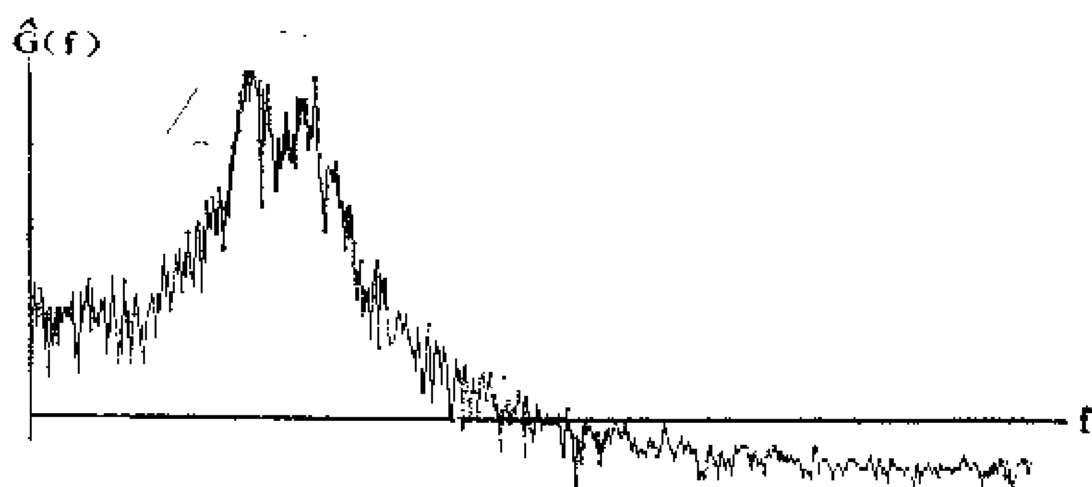
的系统后产生的随机过程，其功率谱的理论曲线如图 4-20*a*）所示，*b*)图是对1024个数据用直接法加工矩形窗但无分段平均时的谱估计曲线，*c*)图是加帕森（*Parzen*)-2 数据窗但无分段平均，*d*)图是加帕森-2数据窗且分段长度为64按50%重叠计算的直接法谱估计曲线，*e*）图则先用直接法（帕森-2 数据窗，分段长度128，50%重叠），再经过间接法（相关函数截取长度为256，用$\beta=10$的修正凯泽迟后窗）得到的谱估计。



(a)



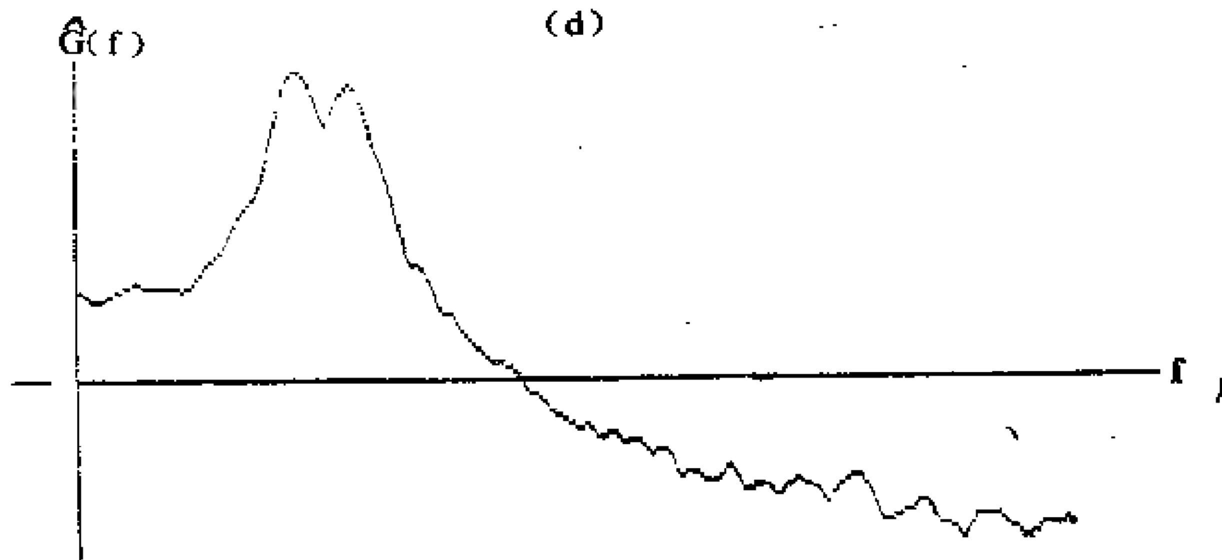
(b)



(c)



(d)



(e)

图4-20 谱估计示例

4·5 富氏变换的细化与高分辨力谱分析

通常由离散富氏变换（如FFT算法）给出的频谱是一系列从零频率开始，等间隔地到某一最高频率f_c的离散谱值（称为谱线）。如果总的点数为 L（即有L条谱线），则相邻谱线间的频率差值为$\triangle_f=f_c/L$。工程上称此为分辨力。分辨力在衡量谱分析的有效性方面是一个很重要的指标。没有适当的频率分辨力，就可能会分不清或遗漏某些应有的频率成份而得出不正确的结论。通常对N点数据作富氏变换所得的频谱有效范围是0到 f_{max}（f_{max}大于或等于信号所含的最高频率成份），而且$f_{max}=1/(2\Delta)$（Δ为时域采样间隔）。各离散频率间的间隔$\Delta_f=1/(N\Delta)=f_{max}/(N/2)=2f_{max}/N$。如当$f_{max}=1000\text{Hz}$，$N=2000$，则$\Delta_f=1$ Hz。也就是说，用2000个采样数据在0到1000Hz范围内作频谱分析只能达到1Hz的分辨力，从图形上看，在0～1000Hz范围内有1000条谱线。

如果只是对上述范围内的某一局部频率区间感兴趣，譬如在900～1000Hz内可能有若干密集的谱峰需要更细致的分辨，那么原有的1000条谱线中实际有用的只有100条。如果有办法在N不变的情况下还能用1000条谱线复盖900～1000Hz的范围，则称其具有十倍的细化功能。

还有一种情况是当信号有很高的频率成份时（f_{max}很大），这时△必需很小或N很大才能满足采样定理，这使分析和处理的时间和费用加大。如果注意到一般方法得到的谱，包括了从0到f_{max}的整个频带，能否利用“频率混叠”效应把高频成份有意识地反映到低频区（即0到f_0，而$f_0<f_{max}$）来观察，从而使谱的实际计算范围缩小呢？回答是肯定的。处理方法基本上是：先选

定所要观察的高频区段，通过带通滤波器（即设定频窗）对数据作频域截取，然后用较低的采样频率（即小于$2f_{max}$的频率）作数据采样并对它作谱估计，这个谱实际上代表原数据序列的高频区段的谱。

上述两种情况的共同实质要求是：使某一局部频率范围的谱，用尽可能多的（受时间、存贮量的限制）谱线表现出来。实现这一目的的富氏变换称细化的富氏变换（*Zoom transform*），也称通带任选的富氏分析BSFA（*Band selectable* Fourier *Analysis*）。

BSFA方法包括带通滤波、数据抽取、FFT计算等三步（见图4-21）。原始数据可以是连续的（相应地可用模拟滤波器），也可以是按采样定理以较高频率（如$2f_{max}$）采得的数据，然后经过数字带通滤波器处理。滤波器带通上下限由需要观察的频率范围选定。

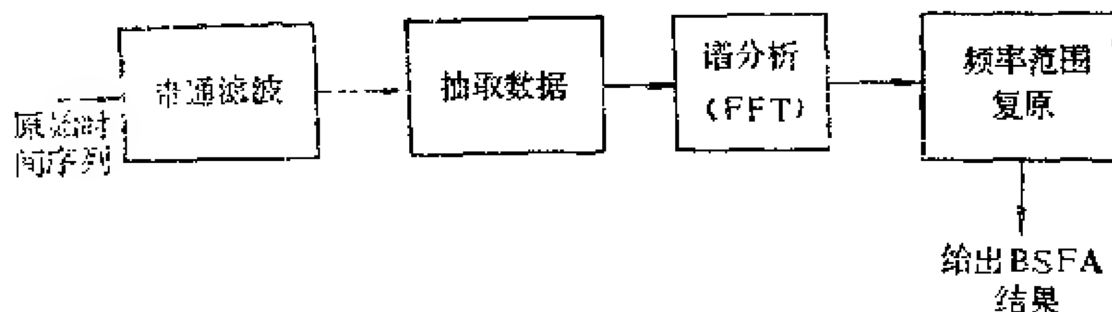


图4-21　细化方法原理图之一

一般将频率范围$0 \sim f_{max} 1/(2\triangle)$分成$P$个区间。通常取$P=2^q$，并称之为$q$阶细化。区间边界为（见图4-22，其中$P=2^3=8$）。

$$\text{左边缘} f_1=\frac{r}{P\cdot 2\triangle}$$

$$r=0,1,\cdots,P-1 \qquad (4\text{-}5\text{-}1)$$

$$\text{右边缘} f_2=\frac{r+1}{P\cdot 2\triangle}$$

选定细化区间也就是确定r，$r=P-1$对应于将频率最高的一段谱作为细化对象。

原始数据通过带通滤波后以较大的间隔$\triangle'$采样

$$\triangle'=P\triangle=2^{q}\triangle \qquad (4\text{-}5\text{-}2)$$

如果原数据是以$\triangle$间隔采样的离散数据，且滤波并未改变其时间

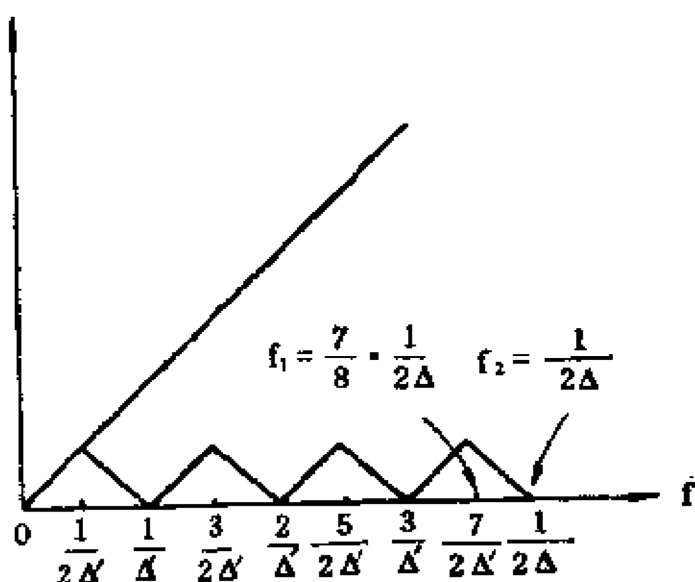


图4-22 细化区间的划分与频率混叠

间隔，则每隔P个数抽取一个，舍弃其余的$P-1$个数。

根据“频率混叠”原理，对经过抽取的数据求谱必然把$f_1\sim f_2$频域内的谱反映到$0\sim 1/2(\triangle')$的范围里来，而由于带通滤波的结果，这一范围内不会有别的频率分量和$f_1\sim f_2$的频率分量混叠。若将反映到0的频率记作f_1'，反映到$1/(2\triangle')=1/(P2\triangle)$的频率记作$f_2'$，则

$$f_1' = \frac{r}{P2\triangle} \qquad f_2' = \frac{r+1}{P2\triangle} \qquad 当 r=0,2,4\cdots,P-2$$

$$f_1' = \frac{r+1}{P2\triangle} \qquad f_2' = \frac{r}{P2\triangle} \qquad 当 r=1,3,5,\cdots P-1 \tag{4-5-3}$$

根据这个关系可以定出细化处理后频率坐标的真值。

在可供处理的数据点数N一定的条件下（通常谱分析仪器中N最大是4096），当采样间隔为△时，分辨力是$1/(N\triangle)$，而$BSFA$的采样间隔取为$P\triangle$，故分辨力可达$1/(NP\triangle)$，提高了P倍，这是细化方法的优点，其分析处理所增加的时间只是带通滤波器所需的时间。

归纳BSFA的步骤如下：

(1)选择分段数P。

(2)建立带通滤波器，其通带上下限为

$$f_{下} = \frac{r}{P2\triangle}, \qquad f_{上} = \frac{r+1}{P2\triangle} \tag{4-5-4}$$

(3)将原序列x_n进行滤波。

(4)对$x_n(n=0,1,2,\cdots)$按$x_n' = x_{(n+1)p}$抽取得新序列x_p，x_{2p}，x_{3p}，…，x_{NP}（N为抽取后的点数）。

(5)利用通常的方法求序列$x_n'(n=0,1,2,\cdots,N-1)$的功率谱$\hat{G}_k(k=0,1,\cdots N-1)$

(6)按以下关系标注正确的频率值

$$k' = \begin{cases} r(N-1)+k & r偶 \\ (r+1)(N-1)-K & r奇 \end{cases} \tag{4-5-5}$$

目前更为常用的另一种细化方法是利用移频—低通滤波技术，它具有较高的细化功能，而且细化区间的选择可以更为灵活，这种方法是首先将密采样的原始序列进行数字移频，即将原序列

x_n乘以$W^{f_0 n}$，其中$W=e^{-j2\pi ZN}$，f_0为细化区间的中心频率值。根据领域位移关系式（见附录式Ⅰ-30），移频前后的频谱应如图4-23a,b所示，将这个移频序列通过带宽等于细化区间的低通滤波器，便可按较低的频率重新抽样得到"稀序列"，并对它作谱分析求得细化频谱，上述方法的原理方框图如图4-24所示。

应用细化富氏变换，使我们能将实际频谱中密集的谱分量扩展开来，便于精确定量分析，由于细化程度愈高需要原始序列愈长，这就要求信号在较长时间内保持平稳，因此分辨力的提高并不能是无限制的。

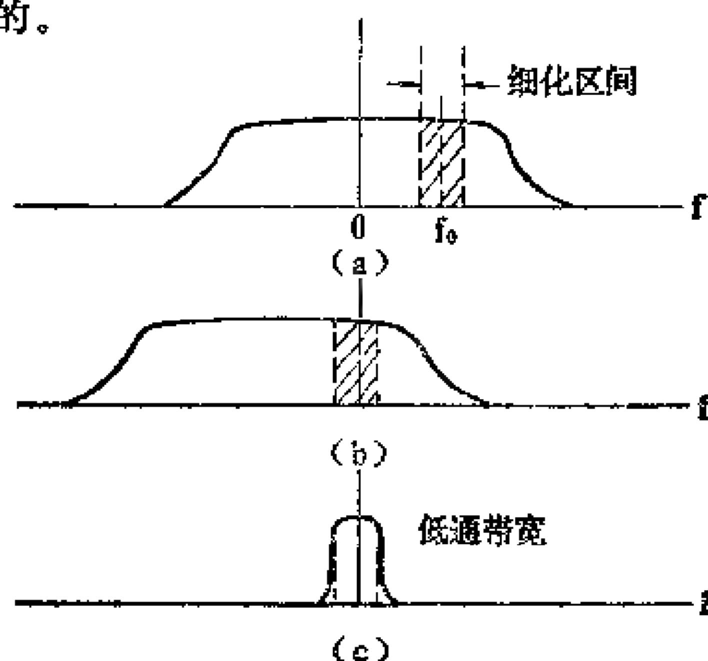
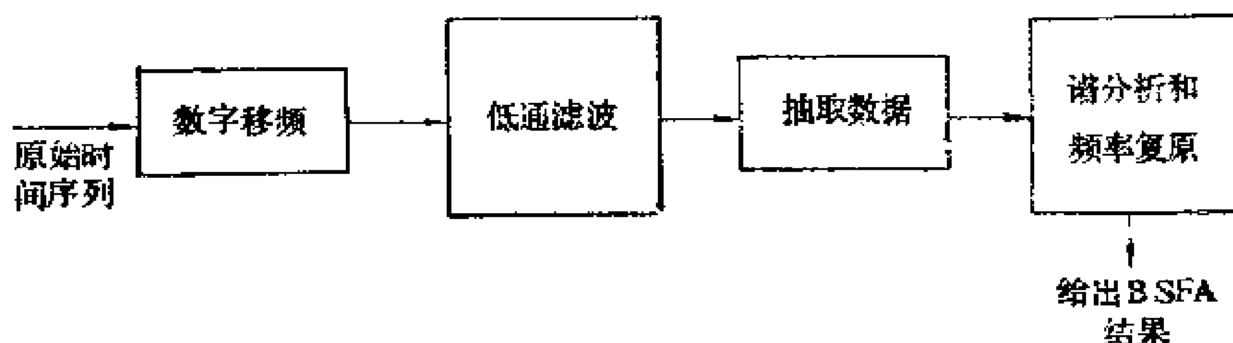


图4-23　频移—低通滤波方法



4-24　细化方法原理图之二

第五章 极大熵谱估计

1967年伯格（J.P.Burg）刚一发表：极大熵谱分析"的方法就在工程和科技界产生很大影响，成为相当流行的功率谱密度估计方法。伯格在谱估计准则的提出和具体算法上有所创新，由此演变出来的算法有很多种，被统称为"现代谱分析"。

5.1 谱熵和极大熵准则

1.问题的提出 从19世纪末舒斯特(Schuster)在利用富氏级数分析信号隐含的周期特性时提出了"周期图"，到1985年由伯来克曼和杜奇提出了谱估计的"间接法"和1965年FFT算法提出后流行的"直接法"，它们本质上都是把原序列经过开窗截取处理来获得对序列谱密度的估计。不论对数据加窗还是对自相关函数加窗，其目的都在于使谱估计的方差减小，然而加窗不可避免地产生频域"泄漏"，使功率谱失真，尽管在窗函数形式的选择和处理方法上做了很多分析研究，使得以周期图为基础的方法达到相当成熟和实用的程度，但是任何抑制旁瓣的方法都是以损失谱分辨力为代价的，这个难题在数据量少的情况下更为突出。

问题的实质是：在周期图估计中，我们对数据或是它的相关函数所做的加窗处理，等于是假定在窗口外数据（或自相关）为零，而窗口内的部分则加上某种形式的修正。这些人为措施使来自观察的信息受到了一定程度的歪曲。

伯格提出的新概念是：和估计的功率谱相对应的自相关和由观察数据算得的自相关一致，同时对已有的区段之外的自相关值采用外推的办法求取，而不是一概假定为零，外推的原则是使相

应的序列在未知点上取值的可能性具有最大的不确定性，亦即不对结果人为地强添任何增加的信息。

数学家申农最早提出“熵”的概念，在统计学中用它作为各种随机试验的不肯定性程度的度量。在热力学和信息论中，“熵”都有其具体的物理背景和应用。伯格将熵引入到谱估计中来提出了极大熵准则的谱估计。后面介绍将会看到，满足熵极大的谱估计是自回归模型的谱。1971年凡登包士（Van Den Bos）证明，一维极大熵谱估计和自回归谱的最小二乘估计是等效的。尽管如此，伯格关于熵谱估计的概念和他对自回归参数的递推算法却独树一帜，随后还有人提出了各种改进算法，但要注意把极大熵概念本身同算法区别开来。

2.高斯过程的熵和熵率 假定我们研究的随机试验α只有有限个不相容的结果A_1，A_2，…，A_n，它们相应的概率为$P(A_1)$，$P(A_2)$，…，$P(A_n)$，且满足$\sum_{i=1}^{n}P(A_i)=1$，简单描述如下

$$\alpha:\begin{pmatrix}A_1, & A_2, & \cdots, & A_n \\ P(A_1), & P(A_2), & \cdots, & P(A_n)\end{pmatrix}$$

申农找到并证明了可以用$H(\alpha)$这个量来度量α的不肯定性的程度：

$$H(\alpha)=-\sum_{i=1}^{n}P(A_i)\log P(A_i)$$

或简写成

$$H(\alpha)=-\sum_{i=1}^{n}P_i\log P_i$$

$H(\alpha)$称为试验α的熵

当随机变量的可能取值是连续的，则H定义式中的和式用积分代替

$$H=-\int_{-\infty}^{\infty}p(x)\log p(x)dx=-E[\log p(x)] \quad (5\text{-}1\text{-}1)$$

其中p(x)为随机变量，对数可以取10或取e为底，在比较熵的大小时并没有影响，下面为计算方便均以自然对数ln来定义，如x为正态随机变量，$p(x)=\frac{1}{\sigma_X\sqrt{2\pi}}e^{-x^2/(2\sigma_X^2)}$，则有

$$H=\ln\sigma_x\sqrt{2\pi e} \quad (5\text{-}1\text{-}2)$$

进一步，如果讨论的是时间序列的实现$\{x_1, x_2, \cdots, x_N\}$则这一过程的熵用下面$N$维积分表示：

$$H=-\int_{-\infty}^{\infty}p(\mathbf{x})\ln p(\mathbf{x})d\mathbf{x} \quad (5\text{-}1\text{-}3)$$

其中$p(\mathbf{x})$是联合概率密度函数$p(x_1, x_2, \cdots, x_N)$，若时间序列是高斯的，则

$$p(\mathbf{x})=[(2\pi)^N\det R_X]^{-1/2}\exp[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_x)_T R_X^{-1}(\mathbf{x}-\boldsymbol{\mu}_x)] \quad (5\text{-}1\text{-}4)$$

其中R_x为自协方差阵

$$R_x\equiv\begin{pmatrix} R_x(0) & R_x(1)\cdots R_x(N-1) \\ R_x(-1) & R_x(0)\cdots R_x(N-2) \\ \cdots\cdots & \\ R_x(1-N) & R_x(2-N)\cdots R_x(0) \end{pmatrix} \quad (5\text{-}1\text{-}5)$$

它的i行j列元素为$R_x(j-i)=E[(x_i-\mu_x)(x_j-\mu_x)]$，$\mu_x$为$x_i$或$x_j$的均值，$\boldsymbol{\mu}_x$表均值向量。

将式(5-1-4)代入式(5-1-3)求过程x的熵

$$H_x=-\int_{-\infty}^{\infty}p(\mathbf{x})\ln[2\pi]^N\det R_x]^{-1/2}d\mathbf{x}+\frac{1}{2}\int_{-\infty}^{\infty}p(\mathbf{x})(\mathbf{x}-\boldsymbol{\mu}_x)^T R_X^{-1}(\mathbf{x}-\boldsymbol{\mu}_x)d\mathbf{x}$$

$$=\frac{1}{2}\ln[(2\pi)^N\det R_x]\int_{-\infty}^{\infty}p(\mathbf{x})d\mathbf{x}$$

$$+\frac{1}{2}\int_{-\infty}^{\infty} trR_x^{-1}(x-\mu_x)(x-\mu_x)^T p(x)dx$$

$$=\frac{1}{2}\ln(\det R_x)+N\ln\sqrt{2\pi}+\frac{1}{2}trR_x^{-1}R_x$$

$$=\frac{1}{2}\ln(\det R_x)+N\ln\sqrt{2\pi e} \tag{5-1-6}$$

式(5-1-6)就是长度为N的正态时间序列的熵。若有正态白噪声ε（方差为σ_ε^2），则

$$p(\varepsilon_1, \cdots, \varepsilon_N)=p(\varepsilon_1)p(\varepsilon_2)\cdots p(\varepsilon_N)$$

$$\ln(\det R_\varepsilon)=N\ln\sigma_\varepsilon^2=2N\ln\sigma_\varepsilon$$

可求得其熵为

$$H_\varepsilon=-E[\ln((p(\varepsilon_1)p(\varepsilon_2)\cdots p(\varepsilon_N))]=N\ln\sigma_\varepsilon\sqrt{2\pi e} \tag{5-1-6}$$

由于H随N增长而发散，定义熵率h为

$$h=\lim_{N\leftarrow\infty}\frac{H}{N} \tag{5-1-8}$$

故白噪声过程的熵率为

$$h_\varepsilon=\ln\sigma_\varepsilon\sqrt{2\pi e} \tag{5-1-9}$$

3.功率谱和熵率的关系　下面给出功率谱和熵率间的一些重要性质和关系。

(1)如果随机向量$y=(y_1, y_2, \cdots y_N)^T$是随机向量$X=(x_1, x_2, \cdots, x_N)^T$的线性变换，即

$$y=Ax \tag{5-1-10}$$

其中A是$N\times N$非奇异矩阵，x的联合概率密度为$p(x_1, x_2, \cdots x_N)$，则由于

$$p(y_1, y_2, \cdots, y_N)=\frac{1}{|A|}p(x_1, x_2, \cdots, x_N) \tag{5-1-11}$$

可得

$$H_y = Hy_1, y_2, \cdots y_N = -E\left[\ln\frac{p(x_1, x_2, \cdots, x_N)}{|A|}\right]$$

$$= -E[\ln p(x_1, x_2, \cdots, x_N) - \ln|A|]$$

$$= Hx_1, x_2, \cdots, x_N + \ln|A|$$

$$= -E[\ln p(x_1, x_2, \cdots, x_N)] + \ln|A| \qquad (5\text{-}1\text{-}12)$$

(2)若x_t是一个稳定的因果系统的输入，该系统的传递函数为$G(B)$（它在单位园内无极点），系统单位脉冲响应为g_t。设x_t在$t=-\infty$时开始输入，因而系统输出y_t是平稳的，以h_x和h_y分别表示x和y的熵率，则

$$h_y = h_x + \frac{1}{2}\int_{-1/2}^{1/2} \ln|G(f)|^2 df \qquad (5\text{-}1\text{-}13)$$

其中

$$G(f) = G(e^{-j2\pi f}) = G(B)_{B=\exp(-j2\pi f)} \qquad (5\text{-}1\text{-}14)$$

证明：若x_t在$t=0$时开始输入，则系统输出为

$$\overline{y_t} = \sum_{i=1}^{t} x_{t-i} g_i \qquad (5\text{-}1\text{-}15)$$

当$t\to\infty$时$\overline{y_t}$趋于平稳的y_t。式(5-1-15)是随机变量$\{x_0, x_1, \cdots x_t\}$通过线性变换Ax成为随机变量$\{\overline{y_0}, \overline{y_1}, \cdots, \overline{y_t}\}$，这里

$$A = \begin{bmatrix} g_0 & 0 & & \\ g_1 & g_0 & 0 & \\ \cdots\cdots & & & \\ g_t & g_{t-1} & \cdots & g_0 \end{bmatrix}$$

$$\det A = g^{t+1}$$

根据式(5-1-12)得

$$H_{\overline{y}} = H_x + (t+1)\ln g_0$$

除以$(t+1)$ 并令$t\to\infty$则

$$h_y=h_x+\ln g_0$$

现在只要证明$\ln g_0$等于式(5-1-13)中的第二项积分就够了。由于$|G(f)|^2=G(f)G(-f)$，在$B=\exp(-j2\pi f)$情况下有

$$\int_{-1/2}^{1/2}\ln|G(f)|^2df=\frac{-1}{2\pi j}\oint_{顺}B^{-1}\ln[G(B)G(B^{-1})]dB$$

这里的线积分是沿单位园进行的，因

$$\oint_{顺}B^{-1}\ln G(B)dB=\oint_{顺}B^{-1}\ln G(B^{-1})dB$$

故式(5-1-13)中的第二项积分等于$\frac{-1}{2\pi j}\oint_{顺}B^{-1}\ln G(B)dB$,所以需要证明

$$\ln g_0=\frac{-1}{2\pi j}\oint_{顺}B^{-1}\ln G(B)dB$$

由于$G(B)$在单位园内是解析的，所以上式中的积分路线可以任意小，当$B\to 0$时$G(B)=g_0$。故上式右边等于$\frac{-1}{2\pi j}\ln g_0\oint_{顺}B^{-1}dB=\frac{-1}{2\pi j}\ln g_0(-2\pi j)=\ln g_0$，证明完成。

(3)若x_t是正态过程，其功率谱$G_x(f)$满足

$$\int_{-1/2}^{1/2}\ln G_x(f)df<\infty \tag{5-1-15}$$

则有

$$h_x=\ln\sqrt{2\pi e}+\frac{1}{2}\int_{-1/2}^{1/2}\ln G_x(f)df \tag{5-1-16}$$

注：$\oint_{顺}$和$\oint_{逆}$分别表示顺时针和逆时针方向的围道积分。

这一结论是不难看出的，因为非白正态过程的功率谱密度$G_x(f)$可以看作是方差为1的白噪声通过频率响应模的平方等于$G_x(f)$的线性系统所产生的过程的谱，因此利用式(5-1-13)和(5-1-9)就可导出式(5-1-16)。

式(5-1-16)给出了过程的功率谱密度和它的熵率之间的关系式，由于右边第一项是常数，比较h_x的大小等价于比较第二项积分的大小，因此称$\int_{-1/2}^{1/2}\ln G_x(f)df$为序列$x_t$的谱熵，并以它作为推导极大熵谱估计的出发点。

例如已知过程x_t的方差为σ_x^2，即

$$\int_{-1/2}^{1/2}G_x(f)df=\sigma_x^2 \tag{5-1-17}$$

要导出能使h_x为最大的功率谱$G_x(f)$。这个问题可以通过求泛函极值来解决。以λ表拉格伦日乘子作泛函

$$\begin{aligned}J(G_x)&=\int_{-1/2}^{1/2}\ln G_x(f)\mathrm{d}f-\lambda\int_{-1/2}^{1/2}G_x(f)df\\&=\int_{-1/2}^{1/2}[\ln G_x(f)-\lambda G_x(f)]df\end{aligned}$$

其变分为

$$\begin{aligned}\delta J(G_x)&=\left[\frac{\partial}{\partial\alpha}J(G_x+\alpha\delta G_x)\right]_{\alpha=0}\\&=\int_{-1/2}^{1/2}\frac{\partial}{\partial\alpha}\Big[\ln(G_x+\alpha\delta G_x)-\lambda(G_x+\\&\qquad\alpha\delta G_x)\Big]_{\alpha=0}df\end{aligned}$$

$$= \int_{-1/2}^{1/2} \left[\frac{1}{G_x(f)} - \lambda \right] \delta G_x(f) df$$

达到极值的条件为$\delta J = 0$，故应有

$$G_x(f) = 1/\lambda$$

代回式(5-1-17)可得$G_x(f) = \sigma_x^2$，即$G_x(f)$必须为常数σ_x^2，因此只有当过程为白噪声时才能使熵率达到最大，这里约束条件是方差为σ_x^2。

5.2 极大熵准则的谱估计

根据伯格所提出的概念，功率谱密度估计的准则应当是：

设$\hat{G}_x(f)$表示估计的谱，则它在满足约束条件

$$\int_{-1/2}^{1/2} \hat{G}_x(f) e^{j2\pi fr} df = R(r) \qquad -M \leqslant r \leqslant M$$

(5-2-1)

的同时，应使谱熵$\int_{-1/2}^{1/2} \ln \hat{G}_x(f) df$达到极大，其中$\hat{G}_x(f)$是$f$的正、实、偶函数，这样对应的$R(r)$自然也是$r$的偶函数。

下面论证满足以上要求的$\hat{G}_x(f)$所应具有的形式。

设已知自相关函数$R(r)$在$-M \leqslant r \leqslant M$内的$2M+1$个值，以$\lambda$表拉格伦日乘子作泛函

$$J(G_x) = \int_{-1/2}^{1/2} \ln G_x(f) df - \sum_{k=-M}^{M} \lambda_k \left[\int_{-1/2}^{1/2} G_x(f) e^{j2\pi kf} df - R(k) \right]$$

$$=\int_{-1/2}^{1/2}\Big\{\ln G_x(f)-\sum_{k=-M}^{M}\lambda_k[G_x(f)e^{j2\pi kf}$$

$$-R(k)]\Big\}df$$

$$\delta J=\int_{-1/2}^{1/2}\frac{\partial}{\partial\alpha}\Big\{\ln(G_x+\alpha\delta G_x)-\sum_{k=-M}^{M}\lambda_k e^{j2\pi kf}(G_x+\alpha$$

$$\delta G_x)\Big\}\Big|_{\alpha=0}df$$

$$=\int_{-1/2}^{1/2}\Big[\frac{1}{G_x(f)}-\sum_{k=-M}^{M}\lambda_k e^{k2\pi kf}\Big]\delta G_x(f)df$$

由$\delta J=0$得

$$\hat{G}_x(f)=1/\sum_{k=-M}^{M}(\lambda_k e^{j2\pi kf}) \qquad (5\text{-}2\text{-}2)$$

这里应有

$$\lambda_{-k}=\lambda_k^* \qquad (5\text{-}2\text{-}3)$$

以保证$\hat{G}_x(f)$是实的。将式(5-2-2)代入式(5-2-1)得

$$R(r)=\int_{-1/2}^{1/2}\frac{e^{k2\pi kf}}{\sum_{k=-M}^{M}\lambda_k e^{j2\pi kf}}df \qquad 0\leqslant r\leqslant M \quad (5\text{-}2\text{-}4)$$

用后移算子$B=e^{j2\pi f}$代替变量f，上式所写成

$$R(r)=\frac{-1}{j2\pi}\oint_{顺}\frac{B^{-r-1}}{\sum_{k=-m}^{m}\lambda_k B^{-k}}dB=\frac{1}{j2\pi}\oint_{逆}\frac{B^{-r-1}}{\sum_{k=-M}^{M}\lambda_k B^{-k}}dB$$

$$0\leqslant r\leqslant M \qquad (5\text{-}2\text{-}5)$$

该积分沿B平面单位园进行，基于式(5-2-3)有

$$\sum_{k=-M}^{M}\lambda_k B^{-k}=\Lambda_M^*(B^{-1})\Lambda_M(B) \tag{5-2-6}$$

其中

$$\Lambda_M^*(B^{-1})=\sum_{k=0}^{M}a_k B^{-k}$$

$$\Lambda_M(B)=\sum_{k=0}^{M}a_k^* B^k$$

不难看出

$$\begin{aligned}\Lambda_M^*(B^{-1})\Lambda_M(B)=&a_M a_0^* B^{-M}+(a_{M-1}a_0^*+a_M a_1^*)B^{-(M-1)}+\cdots\\&+(a_0^* a_0+a_1^* a_1+\cdots+a_M^* a_M)+\cdots\\&+(a_0 a_{M-1}^*+a_1 a_M^*)B^{M-1}+a_0 a_M^* B^M\end{aligned}$$

故

$$\lambda_k=\sum_{i=0}^{M-k}a_i^* a_{i+k}=\lambda_{-k}^*$$

多项式$\Lambda_M(B)$的全部零点均在B平面单位圆外，而$\Lambda_M^*(B^{-1})$的全部零点均在单位圆内，两部分零点是互为倒数分布的。将式(5-2-6)代入式(5-2-5)得

$$R(r)=\frac{1}{j2\pi}\oint_{逆}\frac{B^{-r-1}}{\Lambda_M^*(B^{-1})\Lambda_M(B)}dB \qquad 0\leqslant r\leqslant M$$

构造和式

$$\sum_{k=0}^{M}a_k R^{(r-k)}=\sum_{k=0}^{M}a_k R(-r+k)=\sum_{K=0}^{M}a_K B^{-k}R(-r)$$

$$=\frac{1}{j2\pi}\oint_{逆}\frac{B^{r-1}\sum_{k=0}^{M}a_i B^{-k}}{\Lambda_M^*(B^{-1})\Lambda_M(B)}dB$$

$$=\frac{1}{j2\pi}\oint_{逆}\frac{B^{r-1}}{\Lambda_M(B)}dB \qquad r\geqq 0$$

由于$\Lambda_M(B)$在单位圆内没有零点，上式被积函数除了$r=0$时有极点在原点外，$r\geq 1$时在单位圆内是解析的，根据柯西留数定理可得

$$\sum_{k=0}^{M}\alpha_k R(r-k)=\begin{cases}1/\alpha_0^* & r=0\\ 0 & r=1,2\cdots,M\end{cases}$$

详细写，就是

$$r=0 \quad \alpha_0 R(0)+\alpha_1 R(1)+\cdots+\alpha_M R(M)=1/\alpha_0^*$$

$$r=1 \quad \alpha_0 R(1)+\alpha_1 R(0)+\cdots+\alpha_M R(M-1)=0$$

……

$$r=M \quad \alpha_0 R(M)+\alpha_1 R(M-1)+\cdots+\alpha_M R(0)=0$$

若令

$$\alpha_k/\alpha_0=a_k \qquad k=0,1,\cdots,M$$

$$1/(\alpha_0^*\alpha_0)=P_M \text{或} \quad |\alpha_0|^2=1/P_M \tag{5-2-7}$$

则有

$$\sum_{k=0}^{M}a_k R(r-k)=\begin{cases}P_M & r=0\\ 0 & r=1,2,\cdots,M\end{cases} \tag{5-2-8}$$

$$a_0=1$$

利用式(5-2-7)可将式(5-2-6)写成

$$\sum_{K=-M}^{M}\lambda_k B^{-k}=\frac{1}{P_M}A_M^*(B^{-1})A_M(B) \tag{5-2-9}$$

其中

$$A_M^*(B^{-1})=\sum_{k=0}^{M}a_k B^{-k}$$

$$A_M(B)=\sum_{k=0}^{M}a_k^* B^k$$

将式(5-2-9)代入式(5-2-2)，并考虑到$\hat{G}_x(f)$的偶函数性质，得

$$\hat{G}_x(f)=\frac{P_M}{\left|1+\sum_{k=1}^{M}a_k e^{-j2\pi kf}\right|^2} \tag{5-2-}$$

由结果可以看出，在已知头$M+1$个相关函数值时，以它作为约束条件推出的极大熵谱是AR模型的功率谱。

一般情况下，我们直接观察到的是过程x_t的数据，并不知道相关函数的准确值。因此通常根据$\sum_{t=M+1}^{N}\left[x_t+\sum_{i=1}^{M}a_i x_{t-i}\right]$为最小来求$a_i$，这将得到一组$n$个方程，其形式和式(5-2-8)一样，只是用估计的自相关代替$R(r)$，所以当采用这种估计值时，极大熵法和最小二乘法估计的结果是相同的。

5.3 极大熵谱估计的伯格算法

上面已经指出，由于不确切知道头$M+1$个自相关的真值一般只能用它的估计值代替，在第三章中提到的两种估计算式为

$$\hat{R}_{\mathrm{I}}(r)=\frac{1}{N}\sum_{t=1}^{N-|r|}x_t x_{t+|r|}$$

和

$$\hat{R}_{\mathrm{II}}(r)=\frac{1}{N-|r|}\sum_{t=1}^{N-|r|}x_t x_{t+|r|}$$

$\hat{R}_{\mathrm{I}}$是非负定的，方差较小，但估计偏度随r增加而增大。$\hat{R}$是渐近无偏的，但不能保证非负定性质。因此以上两种算法各有不

足之处。伯格采用的算法是从一阶模型开始逐步增加阶数的递推算法，每步递推都能保证相应的自相关序列是非负定的，而且得到的模型也是平稳的，不仅如此，从后面介绍还可看出，由于采用了双向（正、反）预报误差平方和为最小，提高了数据的利用率，充分挖取数据内含的信息，因此特别有利于短数据的分析和建模。

先看一阶模型

$$x_t + a_1^{(1)} x_{t+1} = \varepsilon_t$$

这里上标是用来标明它属于一阶模型，以下都用它作为递推次数的记号，对平方和

$$\eta'(a_1^{(1)}) = \sum_{t=2}^{N} (x_t + a_1^{(1)} x_{t-1})^2$$

作极小化来选择$a_1^{(1)}$时可能会出现$|a_1^{(1)}| > 1$，从而使模型不平稳，例如序列值为1，2时

$$\eta'(a_1^{(1)}) = (x_2 + a_1^{(1)} x_1)^2 = (2 + a_1^{(1)})^2$$

由 $\quad \partial\eta'(a_0^{(1)})/\partial a_1^{(1)} = 2(2 + a_1^{(1)}) = 0 \quad \Rightarrow a_1^{(1)} = -2$

由于在均方意义下最优的模型参数只取决于序列的自相关而非序列值本身，而序列按反向的时间顺序排列并不改变自相关函数（见2.1节），伯格提出以

$$\eta(a_1^{(1)}) = \sum_{t=2}^{N} [(x_t + a_1^{(1)} x_{t-1})^2 + (x_{t-1} + a_1^{(1)} x_t)^2]$$

作极小化来求$a_1^{(1)}$，上式右边第二个括弧内的$x_{t-1} + a_1^{(1)} x_t$可以看作是由x_t“预报”x_{t-1}的误差，不难看出，由此得到的$a_1^{(1)} = -\sum 2x_t x_{t-1} / \sum (x_t^2 + x_{t-1}^2)$满足$|a_1^{(1)}| < 1$。因此这种将正反双方预报误差的平方和作极小化的办法在一阶的情况下是可行的。如果自然地推广到二阶，则有

$$\eta(a_1^{(2)},\ a_2^{(2)})=\sum_{t=3}^{N}[(x_t+a_1^{(2)}x_{t-1}+a_2^{(2)}x_{t-2})^2+$$

$$x_{t-2}+a_1^{(2)}x_{t-1}+a_2^{(2)}x_t)^2]$$

对它作极小化求$a_1^{(2)}$，$a_2^{(2)}$

然而伯格指出，这样求得的模型参数并不总能满足平稳性条件，但他注意到利文森（*Levinson*）提出的递推算法可以做到这一点，这种算法由$AR(1)$推出$AR(2)$是按以下关系：

$$\begin{pmatrix}1\\a_1^{(2)}\\a_2^{(2)}\end{pmatrix}=\begin{pmatrix}1\\a_1^{(1)}\\0\end{pmatrix}+c^{(2)}\begin{pmatrix}0\\a_1^{(1)}\\1\end{pmatrix}$$

伯格决定不由$\eta(a_1^{(2)},\ a_2^{(2)})$对$a_1^{(2)}$，$a_2^{(2)}$，作极小化，而是把$\eta$看作是$c^{(2)}$的函数，然后对$c^{(2)}$极小化($a_1^{(1)}$已由$\eta(a_1^{(1)}=\min$求得)，为此写出

$$\eta(c^{(2)})=\sum_{t=3}^{N}\Big\{[x_t+(a_1^{(1)}+c^{(2)}a_1^{(1)})x_{t-1}+c^{(2)}x_{t-2}]^2$$

$$+[x_{t-2}+(a_1^{(1)}+c^{(2)}a_1^{(1)})x_{t-1}+c^{(2)}x_t]^2\Big\}$$

$$=\sum_{t=3}^{N}\Big\{[x_t+a_1^{(1)}x_{t-1}+c^{(2)}(a_1^{(1)}x_{t-1}+x_{t-2})]^2+$$

$$[x_{t-2}+a_1^{(1)}x_{t-1}+c^{(2)}(a_1^{(1)}x_{t-1}+x_t]^2\Big\}$$

$$=\sum_{t=3}^{N}\Big\{[\varepsilon_+^{(1)}+c^{(2)}\varepsilon_-^{(1)}]^2+[\varepsilon_-^{(1)}+c^{(2)}\varepsilon_+^{(1)}]^2\Big\}$$

其中

$$\varepsilon_+^{(1)}=x_t+a_1^{(1)}x_{t-1}\qquad \varepsilon_-^{(1)}=x_{t-2}+a_1^{(1)}x_{t-1}$$

$\varepsilon_+^{(1)}$表示一阶模型的正向预报误差，$\varepsilon_-^{(1)}$表示相应的反向预报误差。由$\partial\eta(c^{(2)})/\partial c^{(2)}=0$可得

$$c^{(2)}=-\sum_{t=3}^{N}(2\varepsilon_{+}^{(1)}+\varepsilon_{-}^{(1)})/\sum_{t=3}^{N}(\varepsilon_{+}^{(1)2}+\varepsilon_{-}^{(1)2})$$

因$\Sigma(\varepsilon_{+}^{(1)}+\varepsilon_{-}^{(1)})^2$恒大于零，可以证明$|c^{(2)}|\leqslant 1$。后面将会看到，这对模型的平稳性和自相关函数的非负定性都是很关键的。

求得$c^{(2)}$后，$AR(2)$的参数是

$$\begin{pmatrix}1\\a_1^{(2)}\\a_2^{(2)}\end{pmatrix}=\begin{pmatrix}1\\a_1^{(1)}+c^{(2)}a_1^{(1)}\\c^{(2)}\end{pmatrix}$$

从AR(2)向AR(3)递推的方式和前面类似，令

$$\begin{pmatrix}1\\a_1^{(3)}\\a_2^{(3)}\\a_3^{(3)}\end{pmatrix}=\begin{pmatrix}1\\a_1^{(2)}\\a_2^{(2)}\\0\end{pmatrix}+c^{(3)}\begin{pmatrix}0\\a_2^{(2)}\\a_1^{(1)}\\0\end{pmatrix}$$

于是

$$\begin{aligned}\eta(c^{(3)})=&\sum_{t=3}^{N}\Big\{[x_t+(a_1^{(2)}+c^{(3)}a_2^{(2)})x_{t-1}+(a_2^{(2)}+c^{(3)}\\&a_1^{(2)})x_{t-2}+c^{(3)}x_{t-3}]^2+[x_{t-3}+(a_1^{(2)}+c^{(3)}\\&a_2^{(2)})x_{t-2}+(a_2^{(2)}+c^{(2)}a_1^{(2)})x_{t-1}+c^{(3)}x_t]^2\Big\}\\=&\sum_{t=3}^{N}\Big\{[\varepsilon_{+}^{(2)}+c^{(3)}\varepsilon_{-}^{(2)}]^2+[\varepsilon_{-}^{(2)}+c^{(3)}\varepsilon_{+}^{(2)}]^2\Big\}\end{aligned}$$

其中$\varepsilon_{+}^{(2)}$，$\varepsilon_{-}^{(2)}$分别为二阶模型的正反向预报误差

$$\varepsilon_{+}^{(2)}=x_t+a_1^{(2)}x_{t-1}+a_2^{(2)}x_{t-2}=\varepsilon_{+}^{(1)}+c^{(2)}\varepsilon_{-}^{(1)}$$

$$\varepsilon_{-}^{(2)}=x_{t-3}+a_1^{(2)}x_{t-2}+a_2^{(2)}x_{t-1}=\varepsilon_{-}^{(1)}+c^{(2)}\varepsilon_{+}^{(1)}$$

同样，由$\partial\eta(c^{(3)})/\partial c^{(3)}=0$求得

$$c^{(3)}=-\sum_{t=4}^{N}(2\varepsilon_{+}^{(2)}\varepsilon_{-}^{(2)})/\sum_{t=4}^{N}(\varepsilon_{+}^{(2)2}+\varepsilon_{-}^{(2)2})$$

且有 $|c^{(3)}|\leqslant 1$

这样的递推继续下去，到M阶时的一般算式有

$$
\left.\begin{aligned}
&\varepsilon_+^{(0)}=x_t \qquad \varepsilon_-^{(0)}=x_{t-1} \qquad t=1,2,\cdots,N\\
&c^{(M)}=-\sum_{t=M+2}^{N}2\varepsilon_+^{(M-1)}\varepsilon_-^{(M-1)}\Big/\sum_{t=M+1}^{N}\left[(\varepsilon_+^{(M-1)})^2\right.\\
&\qquad\qquad \left.+(\varepsilon_-^{(M-1)})^2\right]\\
&\varepsilon_+^{(M)}=\varepsilon_+^{(M-1)}+c^{(M)}\varepsilon_-^{(M-1)}\\
&\varepsilon_-^{(M)}=\varepsilon_-^{(M-1)}+c^{(M)}\varepsilon_+^{(M-1)}\\
&a_i^{(M)}=a_i^{(M-1)}+c^{(M)}a_{M-i}^{(M-1)} \qquad i=1,2,\cdots,M-1\\
&a_M^{(M)}=c^{(M)}
\end{aligned}\right\}\quad(5-3-1)
$$

现在来看P_M和自相关函数值的递推式，令

$$
P_0=\hat{R}_x(0)=\frac{1}{N}\sum_{t=1}^{N}x_t^2
$$

阶数从1起$\hat{R}_x(M)$和P_M的递推计算将利用自相关函数阵的对称托普里茨（*Toeplitz*）性质，以二阶向三阶递推为例，由式(5-2-8)写出

$$
\begin{pmatrix}\hat{R}_x(0) & \hat{R}_x(1) & \hat{R}_x(2)\\ \hat{R}_x(1) & \hat{R}_x(0) & \hat{R}_x(1)\\ \hat{R}_x(2) & \hat{R}_x(1) & \hat{R}_x(0)\end{pmatrix}\begin{pmatrix}1\\ a_1^{(2)}\\ a_2^{(2)}\end{pmatrix}=\begin{pmatrix}P_2\\ 0\\ 0\end{pmatrix}\quad(5-3-2)
$$

如果将该矩阵所对应的方程组及变量的顺序反过来，则有

$$
\begin{pmatrix}\hat{R}_x(0) & \hat{R}(1) & \hat{R}_x(2)\\ \hat{R}_x(1) & \hat{R}_x(0) & \hat{R}_x(1)\\ \hat{R}_x(2) & \hat{R}_x(1) & \hat{R}_x(0)\end{pmatrix}\begin{pmatrix}a_2^{(2)}\\ a_1^{(2)}\\ 1\end{pmatrix}=\begin{pmatrix}0\\ 0\\ P_2\end{pmatrix}\quad(5-3-3)
$$

把式(5-3-2)的相关阵扩大到4×4；可得

$$
\begin{bmatrix}
\hat{R}_x(0) & \hat{R}_x(1) & \hat{R}_x(2) & \hat{R}_x(3) \\
\hat{R}_x(1) & \hat{R}_x(0) & \hat{R}_x(1) & \hat{R}_x(2) \\
\hat{R}_x(2) & \hat{R}_x(1) & \hat{R}_x(0) & \hat{R}_x(1) \\
\hat{R}_x(3) & \hat{R}_x(2) & \hat{R}_x(1) & \hat{R}_x(0)
\end{bmatrix}
\begin{bmatrix} 1 \\ a_1^{(2)} \\ a_2^{(2)} \\ 0 \end{bmatrix}
=
\begin{bmatrix} P_2 \\ 0 \\ 0 \\ \triangle^{(3)} \end{bmatrix}
\tag{5-3-4}
$$

其中

$$
\triangle^{(3)} = \hat{R}_x(3) + a_1^{(2)}\hat{R}_x(2) + a_2^{(2)}\hat{R}_x(1) \tag{5-3-5}
$$

对式(5-3-3)也作类似扩大，然后将两个4×4阵按下面方式组合

$$
\begin{bmatrix}
\hat{R}_x(0) & \hat{R}_x(1) & \hat{R}_x(2) & \hat{R}_x(3) \\
\hat{R}_x(1) & \hat{R}_x(0) & \hat{R}_x(1) & \hat{R}_x(2) \\
\hat{R}_x(2) & \hat{R}_x(1) & \hat{R}_x(0) & \hat{R}_x(1) \\
\hat{R}_x(3) & \hat{R}_x(2) & \hat{R}_x(1) & \hat{R}_x(0)
\end{bmatrix}
\left(
\begin{bmatrix} 1 \\ a_1^{(2)} \\ a_2^{(2)} \\ 0 \end{bmatrix}
+ c^{(3)}
\begin{bmatrix} 0 \\ a_2^{(2)} \\ a_2^{(1)} \\ 1 \end{bmatrix}
\right)
$$

$$
=
\begin{bmatrix} P_2 \\ 0 \\ 0 \\ \triangle^{(3)} \end{bmatrix}
+ c^{(3)}
\begin{bmatrix} \triangle^{(3)} \\ 0 \\ 0 \\ P_2 \end{bmatrix}
\tag{5-3-6}
$$

根据式(5-3-1)中关于自回归参数的递推关系可见上式左边有关参数的矩阵乃是三阶模型的参数$[1\ a_1^{(3)}\ a_2^{(3)}\ a_3^{(3)}]^T$，因此上式右边等于$[P_3 0\ 0\ 0]^T$，于是有

$$
\begin{cases}
P_3 = P_2 + c^{(3)}\triangle^{(3)} \\
0 = \triangle^{(3)} + c^{(3)}P_2
\end{cases}
\tag{5-3-7}
$$

解之得　$P_3 = (1 - c^{(3)2})P_2$

利用式(5-3-5)和式(5-3-7)得

$$
\hat{R}_x(3) = -c^{(3)}P_2 - a_1^{(2)}\hat{R}_x(2) - a_2^{(2)}\hat{R}_x(1)
$$

推广到一般，可综合如下递推算式：

$$P_0 = \hat{R}_x(0) = \frac{1}{N}\sum_{i=1}^{N} x_i^2 \qquad (5\text{-}3\text{-}8)$$

$$P_M = P_{M-1}(1 - c^{(M)2}) \qquad (5\text{-}3\text{-}9)$$

$$\hat{R}_x(M) = -c^{(M)}P_{M-1} - \sum_{i=1}^{M-1} a_i^{(M-1)}\hat{R}_x(M-i) \quad (5\text{-}3\text{-}10)$$

以上三式和(5-3-1) 诸式组成了一套极大熵谱估计的伯格递推算法（程序见附录二*MEBURG*）。

最后来说明两个问题

(1)递推所得的参数满足平稳性条件，即

$$A_M(B) = 1 + a_1^{(M)}B + \cdots + a_M^{(M)}B^M = 0$$

的根全部在*B*平面单位圆以外，或者等效地说

$$A_M(B) = (1 - \lambda_1 B)(1 - \lambda_2 B)\cdots(1 - \lambda_M B) = 0$$

中任一λ_i都满足$|\lambda_i| < 1$。

证， 令$B = e^{-i2\pi f}$并将λ_i表为$\lambda_i = |\lambda_i| e^{i2\pi f_i}$，则$\lambda_i B$应当是一个半径小于1的圆，或者说$(1-\lambda_i B)$是圆心在(1,0)但不包围原点的圆(见图5-1)。这必等价于 当*f*由-1/2时变到1/2时$(1-\lambda_i B)$

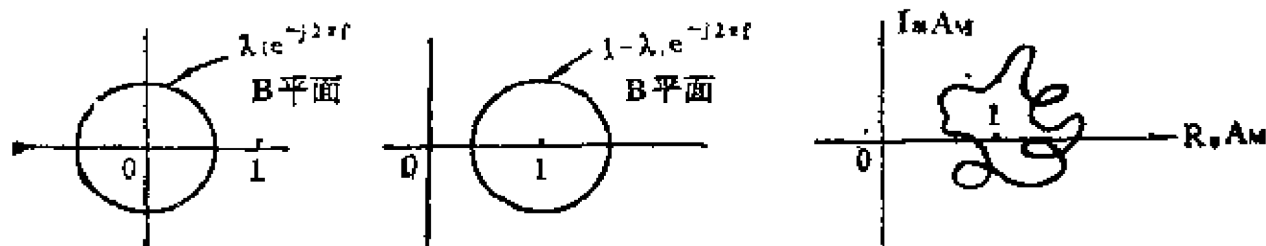


图5-1 A_M (B) 的根满足平稳条件时的曲线

的幅角增量为零。当全部λ_i的模均小于1时，$A_M(B)|_{B=\exp(-i2\pi f)}$总的幅角增量也是零，或$A_M(B)|_{B=\exp(-i2\pi f)}$曲线不包围原点。而今

$$
\begin{aligned}
A_M(B) &= 1 + a_1^{(M)}B + \cdots + a_M^{(M)}B^M \\
&= (1 + a_1^{(M-1)}B + \cdots + a_{M-1}^{(M-1)}B^{M-1}) + c^{(M)}(a_{M-1}^{(M-1)}B \\
&\quad + \cdots + a_1^{(M-1)}B^{M-1} + B^M) \\
&= (1 + a_1^{(M-1)}B + \cdots + a_{M-1}^{(M-1)}B^{M-1})\Bigg[1 + c^{(M)}B^M \\
&\quad \frac{a_{M-1}^{(M-1)}B^{-(M-1)} + \cdots + a^{(M-1)}B^{-1} + 1}{1 + a_1^{(M-1)}B + \cdots + a_{M-1}^{(M-1)}B^{M-1}}\Bigg] \\
&= A_{M-1}(B)\left[1 + c^{(M)}B^M \frac{A_{M-1}(B^{-1})}{A_{M-1}(B)}\right] \qquad (5\text{-}3\text{-}11)
\end{aligned}
$$

这里模

$$
\left| B^M \frac{A_{M-1}(B^{-1})}{A_{M-1}(B)} \right|_{B=\exp(-j2\pi f)} = 1
$$

故当$|c^{(M)}|<1$时

$$
1 + c^{(M)}B^M \frac{A_{M-1}(B^{-1})}{A_{M-1}(B)} \Bigg|_{B=\exp(-j2\pi f)}
$$

是不包围原点的，即式(5-3-11)右边[·]部分的零点都在B平面单位圆外，如果前一步递推得到的$A_{M-1}(B)$已满足平稳条件，则$A_M(B)$也将满足平稳条件。由于从一阶开始递推时已有$|a_1^{(1)}|<1$，且以后每步递推均有$|c^{(i)}|<1$，因此每步递推所得的参数必然均能满足平稳条件。

(2)递推所得的自相关序列满足非负定条件。

证：由于$c^{(M)}\leqslant 1$，根据式(5-3-9)必有$0\leqslant P_M\leqslant P_{M-1}$。再以$\hat{\mathbf{R}}_x^{(M)}$表示由$\hat{R}_x(0)$，$\hat{R}_x(1)$，…，$\hat{R}_x(M)$构成的相关函数阵，则式(5-2-8)可写成

$$
\hat{R}_x^{(M)} \begin{pmatrix} 1 \\ a_1^{(M)} \\ \vdots \\ a_M^{(M)} \end{pmatrix} = \begin{pmatrix} P_M \\ 0 \\ \vdots \\ 0 \end{pmatrix}
$$

由克莱姆法则知

$$\det\hat{R}_x^{(M)}=P_M\det\hat{R}_x^{(M-1)}$$

由于$\det\hat{R}_x^{(0)}=\hat{R}_x(0)=P_0>0$，而$1-c^{(0)2}\geqslant0$，故$P_1\geq0$，同样有$P_2\geq0$，…，按归纳法可知每次递推所得的$\det\hat{R}_x^{(M)}>0$，因此上述递推算法得出的$\hat{R}_x(0)$，$\hat{R}_x(1)$，…，$\hat{R}_x(M)$序列构成正定列。

关于由极大熵谱获得的模型阶数问题，由式(5-2-10)可见，其阶数M是已给自相关估值的最大迟后，当数据个数为N时最大可能的迟后值为$N-1$，这可能并非是过程的真正阶数。而另一方面，如果序列本身是无限阶的AR模型（如$ARMA$模型的等效），需要很高的阶数才能逼近真正的过程，这时已给相关的最大迟后所定出的阶数又可能太小。当然，M愈大，用$\hat{R}_x^{(M)}$估计$R_x^{(M)}$的精度也愈低，所以取很大的阶数未必就好。鉴于极大熵谱是AR谱，我们可以利用诸如FPE、AIC、BIC等定阶准则进行检验和判定。

下面的例子是一组由

$$x_t=3\sin\{0.05[2\pi(t-1)]\}+\varepsilon_t$$

产生的20个数据（$t=1,2,\cdots,20$），其中ε_t是白噪声，它的标准差约为正弦振幅的5%，一个x_t的纪录为0.1410，1.0509，1.7826，2.6804，3.0536，2.9605，2.7524，2.1767，1.6413，1.0371，0.1217，-0.9359，-1.8501，-2.5495，-2.5454，-2.9358，-3.0448，-2.2961，-1.7726，-0.9091（见图5-2），利用MEBURG程序计算可得最佳阶数$n=6$（FPE准则）$a_1=-1.1563$，$a_2=0.5342$，$a_3=-0.6381$，$a_4=0.6302$，$a_5=-0.6210$，$a_6=0.5751$最小残差方差为0.0304。

图5-3为极大熵谱曲线，其峰值出现在0.046Hz处。

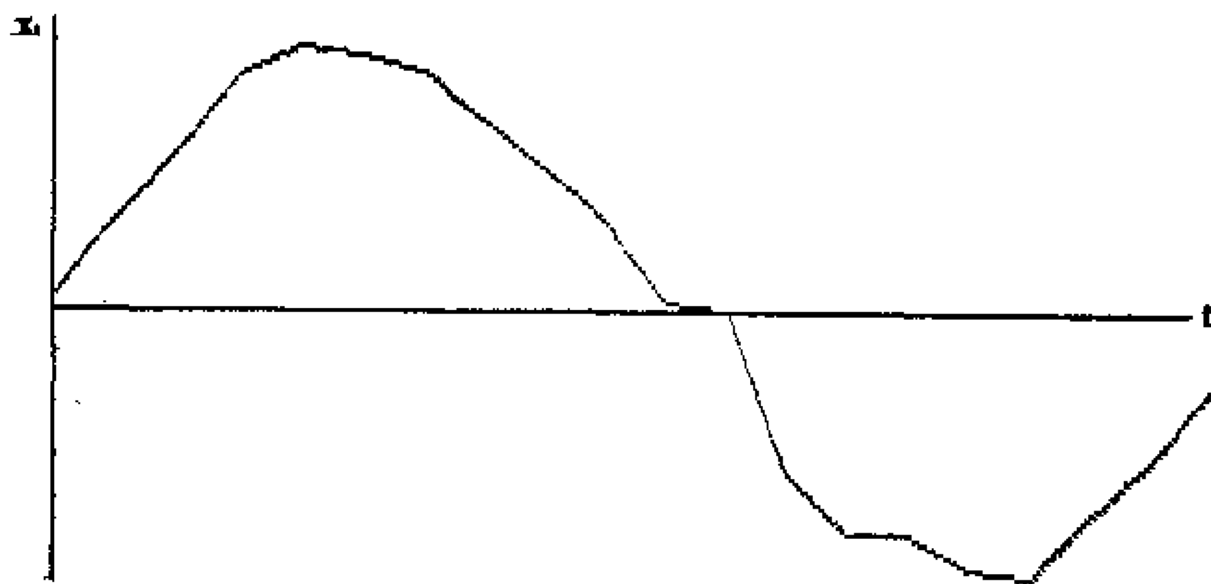


图5－2　20点数据图

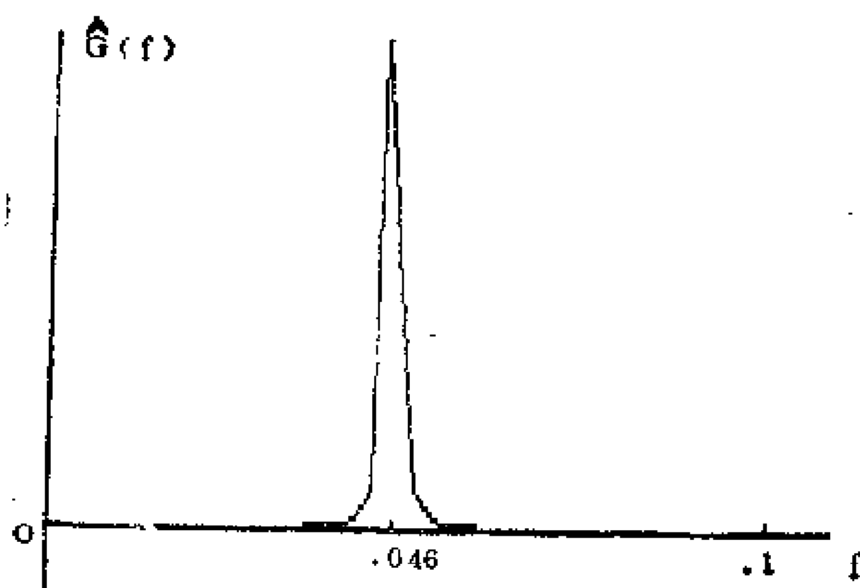


图5-3　20点数据的极大熵谱AR（6）

这里需要指出的是，用自回归模型拟合只适用于有谱密度的序列，对正弦信号而言其谱密度在给定频率处为无穷。这个例子只是说明根据短的序列样本以极大熵谱估计谱线的位置，即正弦

振荡的频率。

极大熵谱估计的伯格算法程序见附录二MEBURG。

5.4极大熵谱估计的LS-LUD算法

伯格在极大熵基础上提出的算法，由于采用了较合理的外推和正反向误差平方和的极小化，比传统的方法提高了分辨力。或者说在同样的分辨力下只需要较少的数据量。在伯格方法引起人们广泛重视和应用的同时，也发现其不足之处，这就是"谱峰偏移"和"谱线分裂"，前者是指峰值频率估值和真值之间的偏离度，后者是指本来只有一个谱峰，但在估计谱中却出现两个或多个相距很近的谱峰。这类现象在耶尔一瓦克尔的AR谱估计中也是存在的(凯伊（*Kay*）和马波（*Marple*）已曾指出过），而伯格算法依然存在这两个问题。泛杰尔(*Fougere*）首先指出：当数据中信噪比高以及所取阶数较高的情况下，伯格算法容易产生谱线分裂，在用周期为T的正弦信号叠加白噪声作样本进行分析中发现，当数据长度为$T/4$的奇数倍，以及正弦的初始相位为$\pi/4$的奇数倍时也容易出现分裂。谱峰偏移也和初始相位有关。而数据量加大，谱线分裂和谱峰偏移量都会减弱。

为了解决上述问题，人们已经并且还在做出努力，通常在研究中都采用正反向预报误差，而实践结果表明，AR模型的最小二乘（LS）解的频率偏移小，在短数据下，不受托普里兹（*Toeplitz*）矩阵形式限制的LS方法可以得到更好的结果。不过LS方法计算量大模，型结果可能非平稳等问题则需要妥善解决。这方面近几年来已取得一些肯定的结果，本节介绍的LS-LUD算法就是在最小二乘方法基础上采用上下三角阵分解（LUD）的算法求解AR谱。

我们知道。若给定时间序列为$x_1, x_2, \cdots, x_N$，当用n个既往观察数据作一步正向预报时。其预报误差为

$$\varepsilon_t^+ \equiv x_t - \hat{x}_t^+ \qquad (5\text{-}4\text{-}1)$$

其中 $$\hat{x}_t^+ = \sum_{i=1}^{n} \varphi_i^+ x_{t-i} \qquad t = n+1, n+2, \cdots, N \qquad (5\text{-}4\text{-}2)$$

根据LS准则，$\sum\limits_{t=n+1}^{N} (\varepsilon_t^+)^2 = \min$，可以确定系数$\varphi_i^+$。

类似地，当用n个后来观察数据作反向预报时。其预报误差记作

$$\varepsilon_t^- \equiv x_t - \hat{x}_t^- \qquad (5\text{-}4\text{-}3)$$

其中 $$\hat{x}_t^- = \sum_{i=1}^{n} \varphi_i^- x_{t+i} \qquad t = 1, 2, \cdots, N-n \qquad (5\text{-}4\text{-}4)$$

而φ_i^-可由$\sum\limits_{t=1}^{N-n} (\varepsilon_t^-)^2 = \min$的准则确定

利用矩阵方程表示时。式(5-4-1)可写成

$$X_+ \Phi_+ - Y_+ = E_+ \qquad (5\text{-}4\text{-}5)$$

其中

$$X_+ \equiv \begin{bmatrix} x_n & x_{n-1} \cdots x_1 \\ x_{n+1} & x_n \quad \cdots x_2 \\ \cdots\cdots & \\ x_{N-1} & x_{N-2} \cdots x_{N-n} \end{bmatrix}$$

$$\Phi_+ \equiv [\varphi_1^+, \varphi_2^+, \cdots \varphi_r^+]^T$$

$$Y_+ \equiv [x_{n+1} x_{n+2} \cdots x_N]^T$$

$$E_+ \equiv [\varepsilon_{n+1}^+ \varepsilon_{n+2}^+ \cdots \varepsilon_N^+]^T$$

式(5-4-5)是含有n个未知参数的$N-n$个方程组。残差平方和记作

$$Q_+ = E_+^{\mathrm{T}} E_+$$

根据$\partial Q_+/\partial \varphi_i^+ = 0$可得$n$个方程式——正规方程：

$$X_+^{\mathrm{T}} X_+ \hat{\Phi}_+ = X_+^{\mathrm{T}} Y_+ \tag{5-4-6}$$

或 $$R_+ \hat{\Phi}_+ = S_+ \tag{5-4-7}$$

其中 $R_+ = [R_{ij}^+] = X_+^{\mathrm{T}} X_+$

$$S_+ = [s_1^+, s_2^+, \cdots, s_n^+]^{\mathrm{T}} = X_+^{\mathrm{T}} Y_+$$

$\hat{\Phi}_+$是Φ_+的LS估计，即$[\hat{\varphi}_1^+, \hat{\varphi}_2^+, \cdots \hat{\varphi}_n^+]^{\mathrm{T}}$

类似地，由反向预报误差可以得出正规方程

$$X_-^{\mathrm{T}} X_- \hat{\Phi}_- = X_-^{\mathrm{T}} Y_- \tag{5-4-8}$$

或 $$R_- \hat{\Phi}_- = S_- \tag{5-4-9}$$

其中 $R_- = [R_{ij}^-] = X_-^{\mathrm{T}} X_-$

$$X_- = \begin{bmatrix} x_2 & x_3 \cdots\cdots x_{n+1} \\ x_3 & x_4 \cdots\cdots x_{n+2} \\ \cdots\cdots & \\ x_{N-n+1} & x_{N-n+3} \cdots x_N \end{bmatrix}$$

$$\hat{\Phi}_- = [\hat{\varphi}_1^- \hat{\varphi}_2^- \cdots \hat{\varphi}_n^-]^{\mathrm{T}}$$

$$Y_- = [x_1 x_2 \cdots x_{N-n}]^{\mathrm{T}}$$

$$S_- = [s_1^- s_2^- \cdots s_n^-]^{\mathrm{T}} = X_-^{\mathrm{T}} Y_-$$

如果同时考虑正反向预报误差，并对两种预报采用同样的自回归系数$\Phi = [\varphi_1 \varphi_2 \cdots \varphi_n]^{\mathrm{T}}$，则可写出含有这$n$个参数的$2(N-n)$个方程式

$$X\Phi - Y = E \tag{5-4-10}$$

其中

$$X = \begin{bmatrix} X_+ \\ X_- \end{bmatrix}, \quad Y = \begin{bmatrix} Y_+ \\ Y_- \end{bmatrix}, \quad E = \begin{bmatrix} E_+ \\ E_- \end{bmatrix}, \quad E_- = [\varepsilon_1^- \varepsilon_2^- \cdots \varepsilon_{N-n}^-]^{\mathrm{T}}$$

而实现 $Q=E^TE=\min$的正规方程为

$$R\hat{\Phi}=S \tag{5-4-11}$$

其中 $R=X^TX=X_+^TX_++X_-^TX_-$

$$S=X^TY=X_+^TY_++X_-^TY_-$$

实际确定参数的过程包含两大步骤，一是正规方程的列写，二是正规方程的求解。一般说，对于含n个未知参数的M个超定方程组，其正规方程的建立需要$Mn^2/2$次运算，用巧列斯基(*Cholesky*)方法求解需要$n^3/6$次运算（一次运算是指一个乘法或除法加上一个加法或减法，在比较运算量时只考虑M和n的最高次幂）。

对应于式(5-4-10)，$M=2(N-n)$，但在求解AR模型参数一特定情况下，由n阶建立$n+1$阶正规方程可以采用递推的方法，下面将表明，对于单向（正向或反向）预报来说，建立正规方程的计算量为$(n^2+5n+2)/2+(N-n-1)$，而对于双向预报的式(5-4-10)，其计算量为$(n^2+5n+2)+(N-n-1)$。由于阶数往往要从1开始搜索，在合适阶数未知的情况下，从1到某个阶数$n2$逐个建立正规方程的总计算量为$N(n+1)+n^3/6+n^2/2-n/3-2$（单向）和$N(n+1)+n^3/3+3n^2/2-5n/6-3$（双向）。

以正向预报中由$n=2$推出$n=3$的情况为例，考虑到正规方程矩阵的对称性，只写它的下三角部分

$$R_+^{(2)}=\begin{bmatrix} R_{11}^{+(2)} & \cdot \\ R_{21}^{+(2)} & R_{22}^{+(2)} \end{bmatrix}=\begin{bmatrix} \sum_{t=1}^{N-2} x_{t+1}x_{t+1} & \cdot \\ \sum_{t=1}^{N-2} x_t x_{t+1} & \sum_{t=1}^{N-2} x_t x_t \end{bmatrix}$$

$$S_+^{(2)}=[s_1^{+(2)}\ s_2^{+(2)}]^{\mathrm{T}}=\left[\sum_{t=1}^{N-2}x_{t+1}x_{t+2}\ \sum_{t=1}^{N-2}x_tx_{t+2}\right]^{\mathrm{T}}$$

$$R_+^{(3)}=\begin{bmatrix} R_{11}^{+(3)} & \cdot & \cdot \\ R_{21}^{+(3)} & R_{22}^{+(3)} & \cdot \\ R_{31}^{+(3)} & R_{32}^{+(3)} & R_{33}^{+(3)} \end{bmatrix}$$

$$=\begin{bmatrix} \sum\limits_{t=1}^{N-3}x_{t+2}x_{t+2} & \cdot & \cdot \\ \sum\limits_{t=1}^{N-3}x_{t+1}x_{t+2} & \sum\limits_{t=1}^{N-3}x_{t+1}x_{t+1} & \cdot \\ \sum\limits_{t=1}^{N-3}x_tx_{t+2} & \sum\limits_{t=1}^{N-3}x_tx_{t+1} & \sum\limits_{t=1}^{N-3}x_tx_t \end{bmatrix}$$

$$S_+^{(3)}=\left[s_1^{+(3)}\ s_2^{+(3)}\ s_3^{+(3)}\right]^{\mathrm{T}}=\left[\sum_{t=1}^{N-3}x_{t+2}x_{t+3}\ \sum_{t=1}^{N-3}x_{t+1}x_{t+3}\ \sum_{t=1}^{N-3}x_tx_{t+3}\right]^{\mathrm{T}}$$

注意到$R_+^{(3)}$中由$1\leqslant j\leqslant i\leqslant 2$构成的$2\times 2$子矩阵的各元素等于$R_+^{(2)}$中的元素减去$R_+^{(2)}$元素所定义的内积中的第一项，即

$$R_{ij}^{+(3)}=R_{ij}^{+(2)}-x_{3-i}x_{3-j}\qquad 1\leqslant j\leqslant i\leqslant 2$$

$R_+^{(3)}$最后一行中除了第一个元素外，都可由$R_+^{(2)}$的最后一行推得，即

$$R_{3j}^{+(3)}=R_{2\,j-1}^{+(2)}-x_{N-2}x_{N+1-j}\qquad 2\leqslant j\leqslant 3$$

$$R_{31}^{+(3)}=s_2^{+(2)}-x_{N-2}x_N$$

此外，$S_+^{(3)}$除了最后一个元素外，亦可由$S_+^{(2)}$推得

$$s_i^{+(3)} = s_i^{+(2)} - x_{3-i}x_3 \qquad 1 \leqslant i \leqslant 2$$

$$s_3^{+(3)} = \sum_{t=1}^{N-3} x_t x_{t+3}$$

可见，除了要算一个内积（元素$S_3^{+(3)}$）外，$R_+^{(3)}$和$S_+^{(3)}$中的其余元素只要一次运算便可求得。

一般情况下，由$R_+^{(n)}$和$S_+^{(n)}$推出$R_+^{(n+1)}$和$S_+^{(n+1)}$时，也只有元素$s_{n+1}^{+(n+1)}$需要多于一次的运算。

关于三种方式下正规方程的建立过程用编程格式列出如下：

(1)正向预报方式算法

$R_{n+1,1}^{+} \longleftarrow s_n^{+} - x_{N-n}x_N$

for $j=2,\cdots,n+1$ *do*

  $R_{n+1,j}^{+} \longleftarrow R_{n,j-1}^{+} - x_{N-n}x_{N+1-j}$

for $i=1,\cdots,n$ *do*

  for $j=1,\cdots,i$ *do*

  $R_{i,j}^{+} \longleftarrow R_{i,j}^{+} - x_{n+1-i}x_{n+1-j}$

for $i=1,\cdots,n$ *do*

  $s_i^{+} \longleftarrow s_i^{+} - x_{n+1-i}x_{n+1}$

$$s_{n+1}^{+} \longleftarrow \sum_{t=1}^{N-n-1} x_t x_{n+1+t}$$

(2)反向预报方式算法

$R_{n+1,1}^{-} \longleftarrow s_n^{-} - x_1 x_{n+1}$

for $j=2,\cdots,n+1$ *do*

  $R_{n+1,j}^{-} \longleftarrow R_{n,j-1}^{-} - x_j x_{n+1}$

for $i=1,\cdots,n$ *do*

  for $j=1,\cdots,i$ *do*

    $R_{i,j}^{-} \longleftarrow R_{i,j}^{-} - x_{N-n+i}x_{N-n+j}$

for $i=1,\cdots,n$ *do*

$s_i^{-} \longleftarrow s_j^{-} - x_{N-n} x_{N-n+i}$

$$s_{n+1}^{-} \longleftarrow \sum_{t=1}^{N-n-1} x_t x_{n+1+t}$$

(3)双向预报方式算法

$R_{n+1,1} \longleftarrow s_n - x_{N-n} x_N - x_1 x_{n+1}$

for $j=2, \cdots, n+1$ *do*

$\quad R_{n+1,j} \longleftarrow R_{n,j-1} - x_{N-n} x_{N+1-j} - x_j x_{n+1}$

for $i=1, \cdots, n$ *do*

$\quad$ *for* $j=1, \cdots i$ *do*

$\quad R_{i,j} \longleftarrow R_{i,j} - x_{n+1-i} x_{n+1-j} - x_{N-n+i} x_{N-n+j}$

for $i-1, \cdots, n$ *do*

$\quad s_i \longleftarrow s_i - x_{n+1-i} x_{n+1} - x_{N-n} x_{N-n+i}$

$$s_{n+1} \longleftarrow 2 \times \sum_{t=1}^{N-n-1} x_n x_{n+1+t}$$

关于$n \times n$正规方程的求解，第一步是将R（它是对称正定阵）分解成LL^T（L为下三角阵），所需运算量为$O(n^3)$，第二步是解$L\Psi=S$得Ψ，然后由$L^T\Phi=\Psi$解得Φ，需要的运算量为$O(n^2)$，两步共需$n^3/6+3n^2/2+O(n)$的运算，这样从$n=1$到某一阶数n独立求解所需的总计算量为$n^4/24+7n^3/12+O(n^2)$。由于计算中采用了三角阵的分解（巧列斯基分解），故简称LUD算法（*Lower and Upper Triangular Matrix Decomposition*）。

总的说来，在$N>>n$情况下，全部计算中占主要的计算是建立正规方程，虽然LS-LUD算法的计算量比伯格算法大，但相差不是数量级的。至于LS-LUD算法可能出现非平稳结果（尽管实际出现的机会很小），可以通过把$A_M(B)$的根映射到单位圆内的办法加以解决。

下面以一个短样本的正弦数据为例，用LS-LUD算法和BU

RG算法作比较，数据是0.03Hz和0.2Hz正弦波的叠加（见图5-4）以每秒10个采样共取75个数据点，因此可供分析的周期数对0.2Hz分量和0.03Hz分量而言分别是1.5周和0.23周。

计算结果的功率谱曲线如图5.5各图所示。（两种算法的曲线形状类似，图a,b,c分别对应于$n=2,3,5$）表5-1列出出现峰值的频率。

表5--1

	AR(2)	AR(3)	AR(4)	AR(4)
伯格算法	0.13HZ	0, 0.199	0.03, 0.202	0.03, 0.202
LUD算法	0.13HZ	0, 0.197	0.03, 0.200	0.03, 0.200

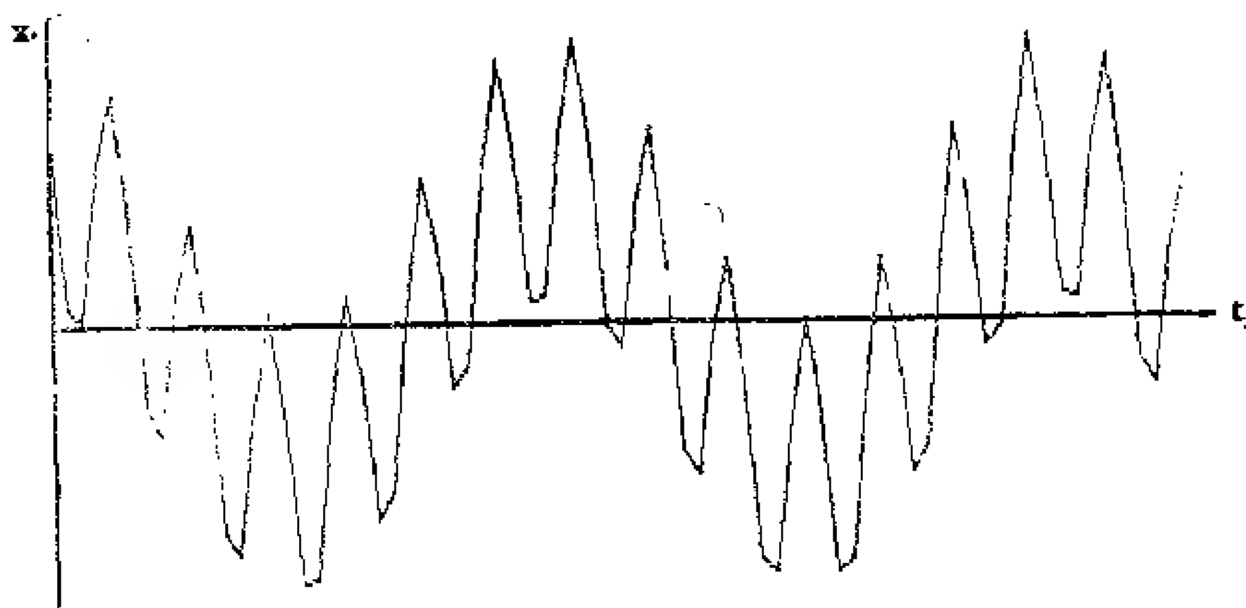


图5-4 频率为0.2和0.03HZ的正弦信号

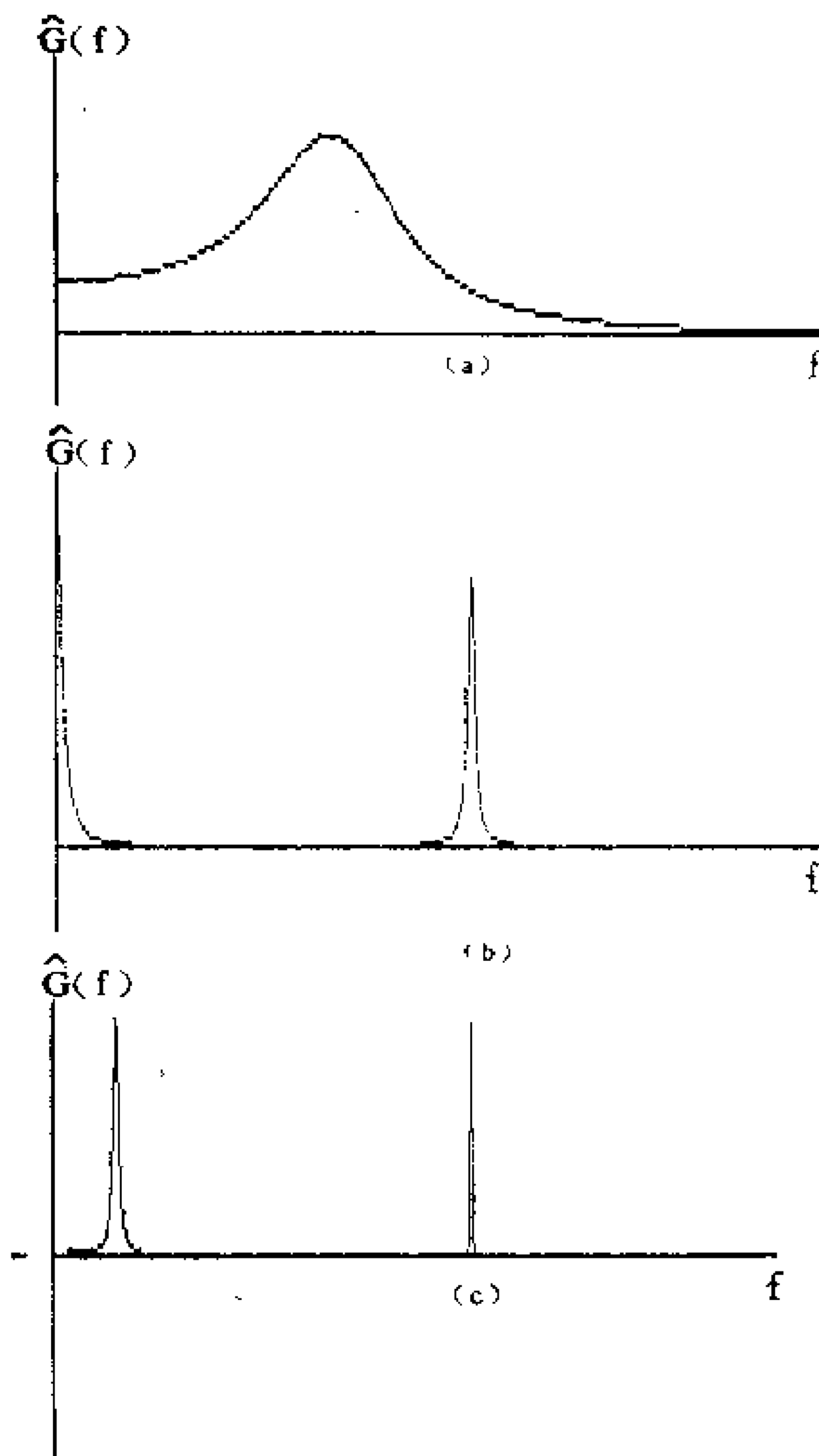


图5-5 两个正弦信号的极大熵谱

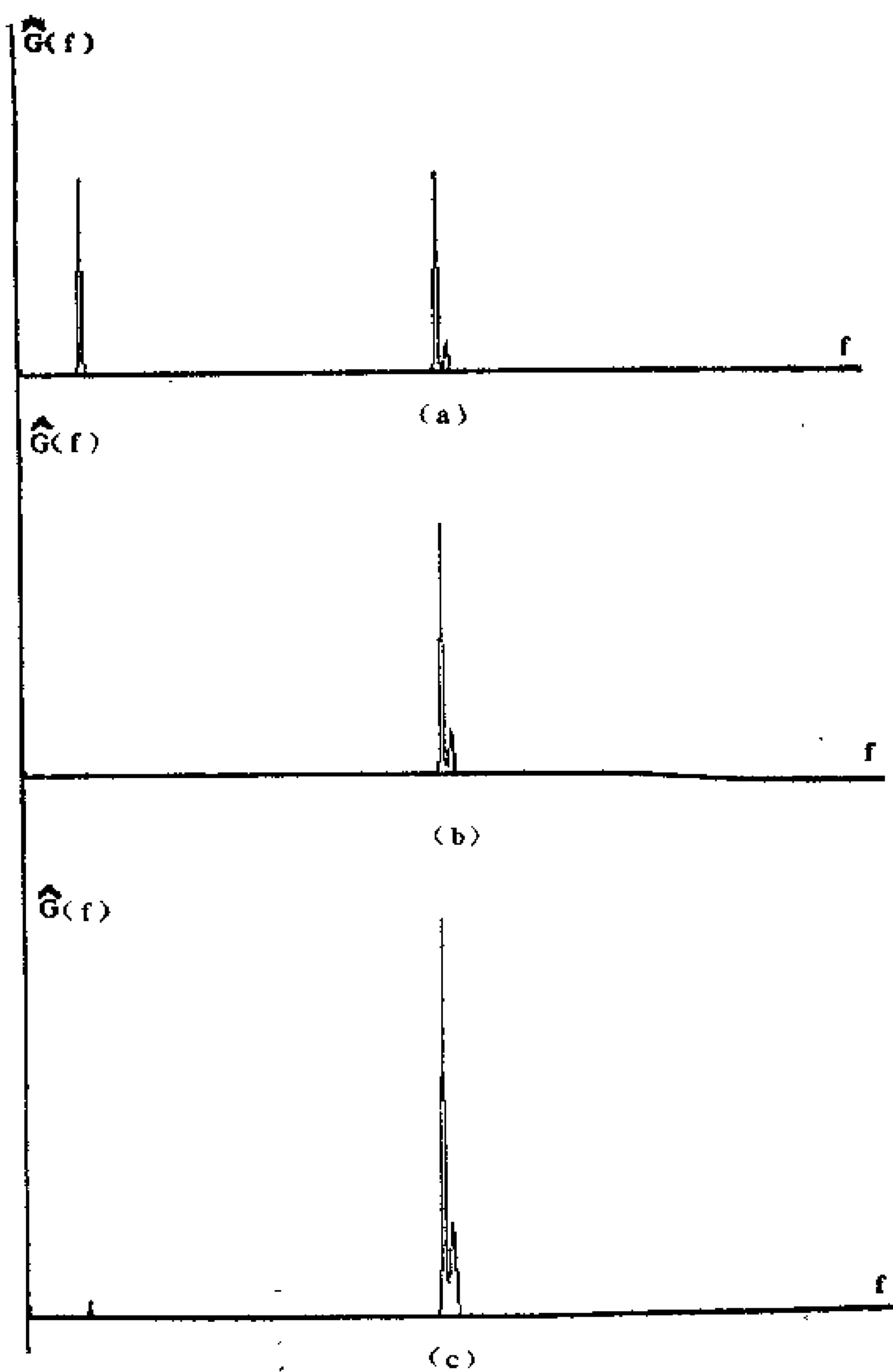


图5-6 谱线分裂示例

可以看出从$n=4$开始LUD算法得出的熵谱已经能够得出真正的频率，而伯格算法对于0.2Hz频率分量有1%的估计偏移，如果进一步提高伯格算法的阶数，不但未能消除这种频率偏移，而且出现谱线分裂，其值如下所列，可以看出这种现象不因n的增加而消失。

AR(7)——0.201，0.206（图5-6a）

AR(10)——0.201，0.205（图5-6b）

AR(12)——0.201，0.205（图5-6c）

应当指出，如果可用的数据周期较多（如15个以上），则伯格算法和LS-LUD算法结果很接近。

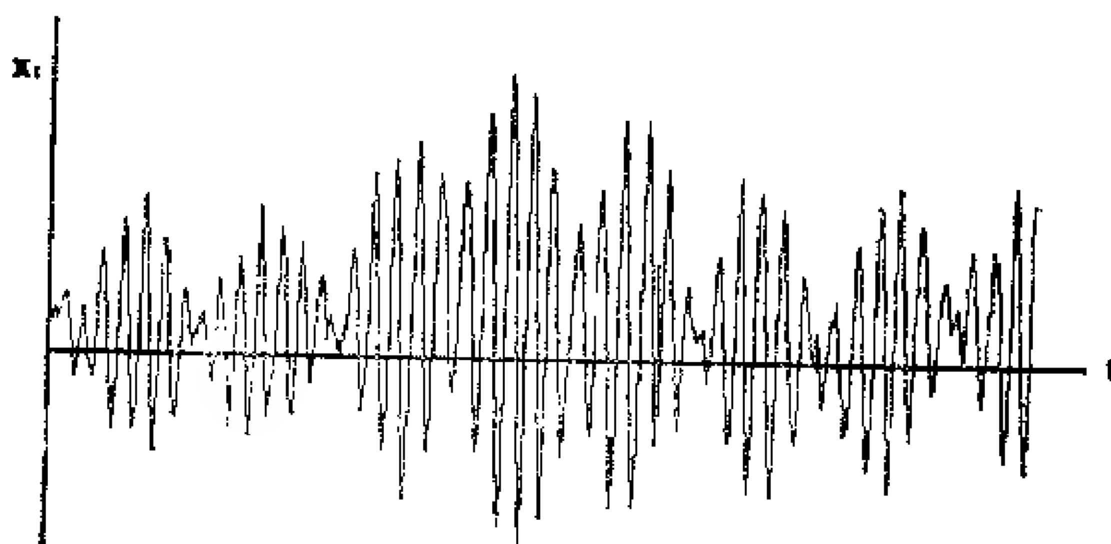


图5—7 512点地极坐标数据图

作为一个实例，我们根据地极坐标资料（见国际极移服务1979年年报）的512个数据，采样间隔为0.1年。（见图5-7）计算的极大熵谱，由LUD算法所得的最佳阶数为$n=27$（残差平方和为0.8676，残差方差为0.8945×10^{-3}），可以看出极移在0.7～1.1年$^{-1}$间有两个周期。（图5-8）

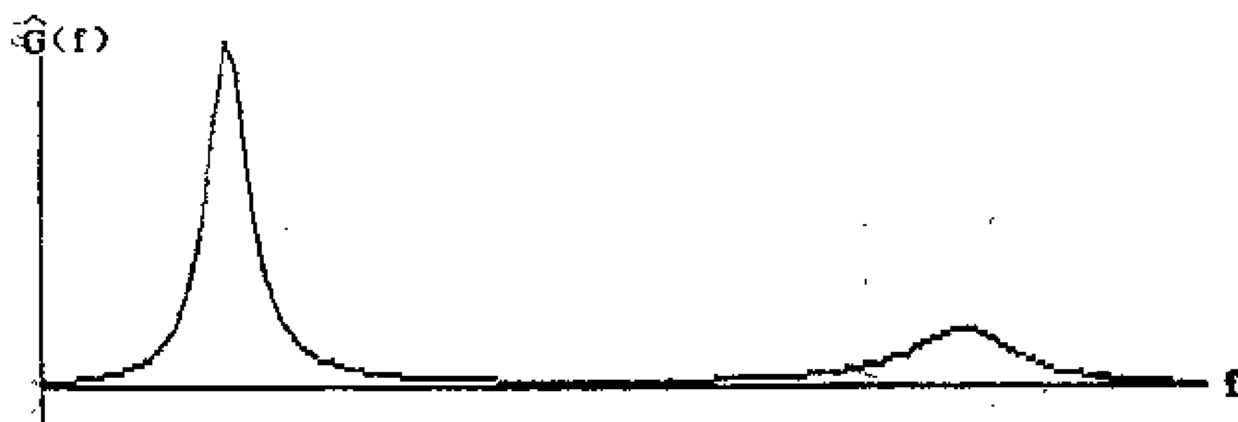


图5-8 地极坐标数据的极大熵谱

第六章　时间序列的预报

在气象、水文、地理、天文等许多实际问题中，我们观察数据进行分析的目的，就是要确定出相应模型及参数，然后对未来可能出现的结果进行预报。这就是说，要根据现在与过去观察序列的样本值，对该序列未来时刻的取值进行估计。既然是进行估计，就要确立衡量估计效果优劣的标准，本章采用的标准是使线性预报的方差达到最小，即是讨论在线性最小方差意义下平稳预报的性质及具体方法。对AR，MA，$ARMA$各类序列预报方法的阐述，着重于它们在计算机上的实现，此外，在本章最后一节将对新息预报方法做简要的介绍。

6.1　平稳线性最小方差预报

1.定义和几何直观解释　设 $\{x_t\}$ 是零均值平稳序列，符号$\hat{x}_t(l)$表示用t时刻及其之前的全部历史观察数据，对未来$t+l$时刻的值x_{t+l}所做的l步平稳线性最小方差预报。

由定义$\hat{x}_t(l)$可表示成如下形式：

$$\hat{x}_t(l)=c_0^lx_t+c_1^lx_{t-1}+c_2^lx_{t-2}+\cdots\cdots \tag{6-1-1}$$

被选定的系数序列 $\{c_i^l\}$ 应使预报方差

$$e_t(l)=x_{t+l}-\hat{x}_t(l)$$

的均方值

$$E[e_t(l)]^2=E[x_{t+l}-\hat{x}_t(l)]^2$$

达到最小。

运用线性空间的观点可以清楚地解释最小方差预报的几何意

义。我们先来回顾一下线性空间和投影的几何意义。

假设有两个向量$\mathbf{x}_1$和$\mathbf{x}_2$，我们要用它们的线性组合来近似表达第三个向量$\mathbf{x}_3$。换句话说，就是说要求出系数a_1和a_2使得向量$a_1\mathbf{x}_1+a_2\mathbf{x}_2$和向量$\mathbf{x}_3$最接近，或者说使它们之间的距离最短。如图6-1所示，这个解可以从向量$\mathbf{x}_3$作垂线到$\mathbf{x}_1$，$\mathbf{x}_2$所张成的平面而得，即$\mathbf{x}_3$的正交投影所给出的向量$\hat{\mathbf{x}}_3=a_1\mathbf{x}_1+a_2\mathbf{x}_2$是最接近于$\mathbf{x}_3$的解。

现在的预报问题是已知x_t，x_{t-1}，x_{t-2}，……，求c_0^t，c_1^t，c_2^t，……使得式(6-1-1)所构成的$\hat{x}_t(l)$和x_{t+l}的距离最近。由于两个随机向量间的距离是用它们差值的均方根表示，因此满足这个条件的解就使$E[e_t(l)]^2$为最小。在维数有限(例如三维)的情况下，从图形上可以直观地看出几何意义。如图6-1中所示，其中$\mathbf{x}_1$，$\mathbf{x}_2$并不一定是垂直的，因此也不相互正交，所以寻求$\hat{x}_3$(或c_1^t，c_2^t)并不方便。如果用一组正交的坐标系，图6-2，其中$\mathbf{e}_1$，$\mathbf{e}_2$为正交单位向量，则可将$\mathbf{x}_3$的正交投影表示成$\hat{\mathbf{x}}_3=a_1'\mathbf{e}_1+a_2'\mathbf{e}_2$。

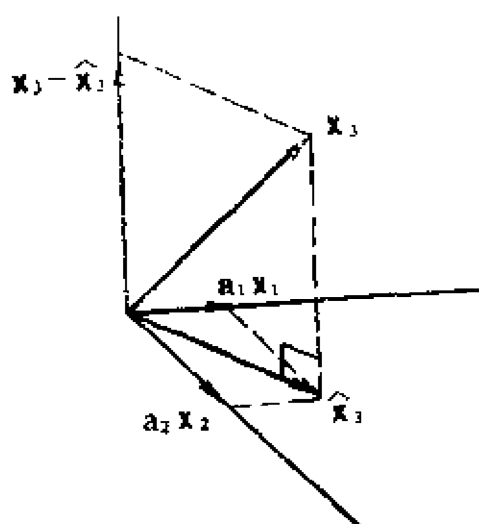


图6-1

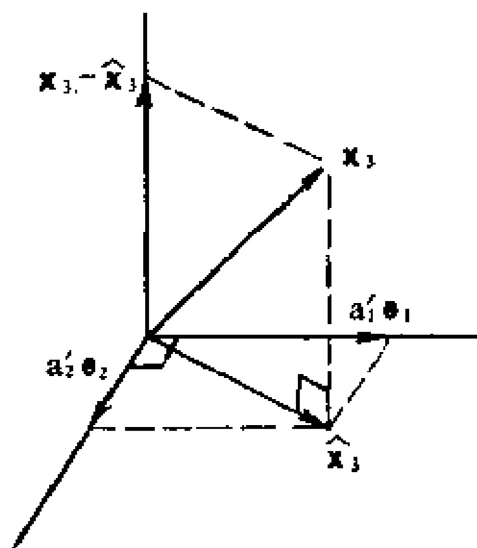


图6-2

对于预报问题，通过引入线性空间及投影的概念，我们可以对预报的实质有形象而直观的了解。为叙述方便，我们引入下面

线性空间。设随机序列所适合的$ARMA$模型是

$$x_t+a_1x_{t-1}+\cdots+a_nx_{t-n}=\varepsilon_t+b_1\varepsilon_{t-1}+\cdots+b_m\varepsilon_{t-m}$$

定义线性空间

$$\mathscr{E}_t=\left\{\varepsilon:\ \varepsilon=\sum_{j=0}^{\infty}d_j\varepsilon_{t-j},d_j\text{为实数},\sum_{j=0}^{\infty}d_j^2<+\infty\right\}$$

$$\mathscr{X}_t=\left\{x:\ x=\sum_{j=0}^{\infty}c_jx_{t-j},c_j\text{为实数},\sum_{j=0}^{\infty}c_j^2<+\infty\right\}$$

$$\mathscr{X}=\{x:\ x\text{是方差有穷随机变量}\}$$

由定义可知$\mathscr{X}_t$是$\mathscr{X}$的子空间，即

$$\mathscr{X}_t=\mathscr{X}$$

由$ARMA$模型具有等价的传递形式和逆转形式，可知实际上

$$\mathscr{X}_t=\mathscr{E}_t$$

也就是说，对任何$y\in\mathscr{E}_t$，必然存在实数u_j，$\sum_{j=0}^{\infty}u_j^2<+\infty$，有$y=\sum_{j=}^{\infty}u_j\varepsilon_{t-j}$，利用$ARMA$的逆转形式，每个$\varepsilon_{t-j}$可表成$x_t$，$0\leqslant t<+\infty$的线性组合，因此必有$y\in X_t$；反之，对任何$y\in\mathscr{X}_t$ 利用传递形式可以导出$y\in\mathscr{E}_t$，这就表明两个线性空间相等。

对于$l>0$，显然x_{t+l}不能用x_t，x_{t-1}，…的线性组合表示，因此x_{t+l}不属于$\mathscr{X}_t$而由定义$x_{t+l}\in\mathscr{X}$，这就是说x_{t+l}属于空间$\mathscr{X}$，但不属于其子空间$\mathscr{X}_t$。

在空间$\mathscr{X}$中，两个元u和v之间的距离用它们差值的均方根表示

$$d(u,\ v)=(E[u-v]^2)^{\frac{1}{2}}$$

线性最小方差预报，就是要选择$x^0\in\mathscr{X}_t$，取$\hat{x}_t(l)=x^0$，在线性

空间$\mathscr{X}_t$中x^0与x_{t+l}的距离较之$\mathscr{X}_t$中其它任何点都短。很自然。地$\hat{x}_t(l)$应取做x_{t+l}在空间X_t上的正交投影。用$\widehat{E}$表示正交投影，则

$$\hat{x}_t(l)=\widehat{E}(x_{t+l}|\mathscr{X}_t)$$

由投影的性质有

$$e_t(l)\perp\mathscr{X}_t$$

符号“⊥”表示垂直(或正交)。由$ARMA$模型白噪声的定义知：$E(\varepsilon_t\varepsilon_{t+\gamma})=0(r\neq0)$，因此$\varepsilon_t$是相互无关或正交的，因而$\varepsilon_{t-j}$，$j=0,1,2,\cdots$是空间$\mathscr{X}_t$中的一组互相正交的座标系。我们可以将$\hat{x}_t(l)$表示成$\varepsilon_{t-j}$，$j=0,1,2,\cdots$的线性组合

$$\hat{x}_t(l)=d'_0\varepsilon_t+d'_1\varepsilon_{t-1}+\cdots+\sum_{j=0}^{\infty}d'_j\varepsilon_{t-j}$$

$\hat{x}_t(l)$要能作为x_{t+l}对$\varepsilon_t,\varepsilon_{t-1},\cdots$所张成空间$\mathscr{X}_t$上的正交投影，必须且只须预报误差$e_t(l)$对每个$\varepsilon_{t-j}(j=0,1,2,\cdots)$正交，即满足下面条件：

$$E[e_t(l)\varepsilon_{t-j}]=E[(x_{t+l}-\hat{x}_t(l))\varepsilon_{t-j}]=0$$
$$j=0,1,2,\cdots$$

将x_{t+l}表示成传递形式，得到

$$E[((\varepsilon_{t+l}+G_1\varepsilon_{t+l-1}+\cdots+G_l\varepsilon_t+G_{l+1}\varepsilon_{t-1}+\cdots)-(d'_0\varepsilon_t+d'_1\varepsilon_{t-1}+\cdots))\varepsilon_{t-j}]=0\qquad j=0,1,2,\cdots$$

由$j=0$可得

$$d'_0=G_l$$

类似地，$j=1$时

$$E[(G_{l+1}-d'_1)\varepsilon^2_{t-1}]=\sigma^2_\varepsilon(G_{l+1}-d'_1)=0$$

故

$$d'_1=G_{l+1}$$

因此实现最小线性方差预报的解可表为

$$d'_j=G_{j+l}\qquad j=0,1,2,\cdots$$

预报误差为

$$e_t(l)=x_{t+l}-\hat{x}_t(l)=\varepsilon_{t+l}+G_1\varepsilon_{t+l-1}+\cdots+G_{l-1}\varepsilon_{t+1} \quad (6\text{-}1\text{-}2)$$

预报误差均方为

$$\mathrm{Var}[e_t(l)]=\sigma_\varepsilon^2(1+G_1^2+\cdots+G_{l-1}^2) \qquad (6\text{-}1\text{-}3)$$

(6-1-3)式表明，l步线性最小方差预报的预报误差方差仅与预报步长l有关，而与预报的起点t无关，这一点体现了预报的平稳性质。由式(6-1-3)还可以看出，预报步长l越大，预报误差方差也越大，即预报准确性越差。

利用$ARMA$模型的传递形式，并参照(6-1-2)可知

$$x_{t+l}=\underbrace{\varepsilon_{t+l}+G_1\varepsilon_{t+l-1}+\cdots+G_{l-1}\varepsilon_{t+1}}_{e_t(l)}+\underbrace{G_l\varepsilon_t+\cdots}_{\hat{x}_t(l)}$$

其中前l项是l步预报误差，后面其余各项是l步线性最小方差预报值：

$$\hat{x}_t(l)=\sum_{j=0}^{\infty}G_{l+j}\varepsilon_{t-j} \qquad (6\text{-}1\text{-}4)$$

格林函数G_j可由$ARMA$模型的参数递推计算出。式(6-1-4)给出了计算l步预报值$\hat{x}_t(l)$的格林函数方法。

综合上述我们已经阐明，线性最小方差预报$\hat{x}(l)$就是x_{t+l}在线性空间$\mathscr{H}_t$上的投影。对于过程$\{x_t\}$为正态的情形，还可以证明[4]$\hat{x}_t(l)$恰恰就是在给定$\mathscr{H}_t$时x_{t+l}的条件期望。也就是在正态条件下，线性最小方差预报、正交投影、条件期望这三者是完全等价的。

2. 最小方差预报的性质 根据线性最小方差预报的定义及正交投影、条件期望的基本特性，我们对平稳最小方差预报做进一步的讨论，可得出一系列简单性质。

$$(1)\ \hat{E}\left[\sum_{j=0}^{\infty}c_j x_{t+j}\,\middle|\,\mathscr{H}_t\right]=\sum_{j=0}^{\infty}c_j\hat{E}[x_{t+j}\mid\mathscr{H}_t] \quad (6\text{-}1\text{-}5)$$

该式由正交投影的线性运算直接可得，说明最小方差预报具有线性性质。

(2) $\hat{E}[x_{t+l}|\mathscr{X}_t]=\hat{E}[x_{t+l}|\varepsilon_t]=\begin{cases}\hat{x}_t(l) & l>0\\ x_{t+l} & l\leqslant 0\end{cases}$ (6-1-6)

这里利用了6.1.1中说明过的$\mathscr{X}_t=\mathscr{E}_t$，注意当$l\leqslant 0$时，$x_{t+l}\in\mathscr{X}_t$，故$\hat{x}_t(l)=\hat{E}[x_{t+l}|\mathscr{X}_t]=x_{t+l}$。该性质表明对现在和过去观察值的预报就是它们自身；对未来观察值的预报可通过正交投影得到。

(3) $\hat{E}[\varepsilon_{t+l}|\mathscr{E}_t]=\hat{E}[\varepsilon_{t+l}|\mathscr{X}_t]=\begin{cases}0 & l>0\\ \varepsilon_{t+l} & l\leqslant 0\end{cases}$ (6-1-7)

当$l>0$时$\varepsilon_{t+l}\perp\varepsilon_t$，这可由$E[\varepsilon_i\varepsilon_j]=0$ $(i\neq j)$得到。该式表明对现在和过去白噪声“冲击”的预报，就是其自身，对未来冲量的预报，就是冲量的平均值——零。

从以上三个基本性质出发，不难得出下列性质：

(4)对$ARMA(n,m)$模型，可证明m步以后的预报

$$\hat{x}_t(l)=-a_1\hat{x}_t(l-1)-a_2\hat{x}_t(l-2)-\cdots-a_n\hat{x}_t(l-n) \quad (6\text{-}1\text{-}8)$$

$$l>m$$

该式表明，若m步以内的预报值$\hat{x}_t(1),\hat{x}_t(2),\cdots,\hat{x}_t(m)$已经得到，超过$m$步的预报值可由式(6-1-8)递推得出。

特别，当$\{x_t\}$是$MA(m)$序列，式(6-1-8)成为

$$\hat{x}_t(l)=0 \qquad l>m$$

这意味着对于m阶滑动平均序列，超过m步的预报值恒等于零。

(5) 由(1)—(3)及逆转形式$x_t=\sum\limits_{j=1}^{\infty}I_j x_{t-j}+\varepsilon_t$可推得

$$\hat{x}_t(l)=\sum_{j=1}^{l-1}I_j\hat{x}_t(l-j)+\sum_{j=l}^{\infty}I_j x_{t+l-j} \qquad (6\text{-}1\text{-}9)$$

由于$\hat{x}_t(l-j)$，$j=1,2,\cdots,l-1$ 也是$x_{t-j}(j=0,1,2,\cdots)$的线性组合，所以可令

$$\hat{x}_t(l)=\sum_{j=1}^{\infty}I_j^{(l)}x_{t+1-j}=I^{(l)}(B)x_t \qquad (6\text{-}1\text{-}10)$$

式中$I^{(l)}(B)=\sum_{j=0}^{\infty}I_j^{(l)}B^{j-1}$将(6-1-10)代入(6-1-9)，比较两边系数便可得到

$$\begin{cases} I_j^{(l)}=I_{j+l-1}+\sum_{i=1}^{l-1}I_iI_j^{(l-i)} & l>1 \\ I_j^{(1)}=I_j \end{cases} \qquad (6\text{-}1\text{-}11)$$

由逆函数的递推公式(2,3节式2-3-79)及(6-1-10),(6-1-11)，可直接从已知的观察数据出发，对$ARMA(n,m)$序列的未来值实现线性平稳最小方差预报，我们称之为平稳预报的逆函数方法。

(6)一步预报误差方差就是σ_ε^2。

该性质由式(6-1-2)令$l=1$得：

$$e_t(1)=x_{t+1}-\hat{x}_t(1)=\varepsilon_{t+1} \qquad (6\text{-}1\text{-}12)$$

可见一步预报误差就是残差，因此$\mathrm{Var}[e_t(1)]=\sigma_\varepsilon^2$。

(7)$\hat{x}_{t+1}(l)=\hat{x}_t(l+1)-G_1\hat{x}_t(1)+G_1x_{t+1}$ (6-1-13)

其证明如下：由式(6-1-4)可以写出

$$\hat{x}_{t+1}(l)=\sum_{j=0}^{\infty}G_{j+l}\varepsilon_{t+1-j}$$

和

$$\hat{x}_t(l+1)=\sum_{j=0}^{\infty}G_{j+l+1}\varepsilon_{t-j}$$

两式相减得 $\hat{x}_{t+1}(l)=\hat{x}_t(l+1)+G_l\varepsilon_{t+1}$ 再利用式(6-1-12)便得式(6-1-13)。

式(6-1-13)的含义在于，如果我们把$t+1$时刻的观察值及预报值称为“新”的，而把t时刻的预报值称为“旧”的，那么式(6-1-13)表明：“新”的预报值可由“旧”的预报值及“新”的观察值推算出，我们称该式为递推预报公式。

根据以上性质和公式，我们可以对各种类模型的平稳序列进行线性平稳最小方差预报。

在正态假定下，给定在$x_t, x_{t-1}, \cdots$时，x_{t+l}的条件分布服从$\mathcal{N}(\hat{x}_t(l), \mathrm{Var}[e_t(l)])$，因此，在得到$\hat{x}_t(l)$之后，可以给出$x_{t+l}$的置信区间：

95%的置信上、下限是$\hat{x}_t(l) \pm 1.96(1+G_1^2+\cdots+G_{l-1}^2)^{\frac{1}{2}}$

90%的置信上、下限是$\hat{x}_t(l) \pm 1.65(1+G_1^2+\cdots+G_{l-1}^2)^{\frac{1}{2}}$

6.2. AR及MA序列的预报方法

先讨论AR序列的预报方法，在式(6-1-8)中令$m=0$可得到AR(n)序列的递推预报公式

$$\hat{x}_t(l)=\varphi_1\hat{x}_t(l-1)+\varphi_2\hat{x}_t(l-2)+\cdots+\varphi_n\hat{x}_t(l-n) \qquad l>0 \qquad (6\text{-}2\text{-}1)$$

由于$\hat{x}_t(j)=x_{t-j}$(当$j\leqslant 0$)，令$\Phi(B)=\sum_{i=0}^{n}a_iB^i$，$a_0=1$，式(6-2-1)写成

$$\Phi(B)\hat{x}_t(l)=0 \qquad l>0$$

这是$\hat{x}_t(l)$关于j的n阶差分方程，初值为$\hat{x}_t(0)=x_t$，$\hat{x}_t(-1)=x_{t-1}, \cdots, \hat{x}_t(-n+1)=x_{t-n+1}$，该差分方程可用递推方法直接求解。

由上述可知，对于AR(n)序列而言，只要知道$x_t, x_{t-1}, \cdots, x_{t-n+1}$这几个历史观察值，无需更早的历史资料$x_j$($j\leqslant t-n$)，只要存贮几个值就可以严格地算出该序列的任意步平稳最小方差预报。

MA序列的预报比AR序列复杂。由(6-1-10)，(6-1-11)二式可知对MA序列有

$$\hat{x}_t(l)=\begin{cases}\sum\limits_{j=0}^{\infty} I_j^{(l)} x_{t+1-j} & 0<l\leqslant m \\ 0 & l>m\end{cases} \qquad (6\text{-}2\text{-}2)$$

对于MA(m)序列而言，虽然超过m步的预报值恒为零，但由式(6-2-2)可以看出，当$1\leqslant l\leqslant m$时，预报值$\hat{x}_t(l)$依赖于全部历史数据x_{t-j}，$j\geqslant 0$，这是与AR序列预报的根本区别。当然实际观察数据总是有限的，式(6-2-2)中的无穷和若用有穷和代替，我们得到近似的平稳最小方差预报

$$\hat{x}_t(l)\cong\sum_{j=1}^{K} I_j^{(l)} x_{t+1-j} \qquad 0<l\leqslant m$$

$K\leqslant t$是选定的充分大的数。如果MA序列是可逆的，I_j便具有负指数收敛性，因而$I_j^{(l)}$也是负指数衰减的，我们可根据$I_j^{(l)}$的收敛速度选择适当的K，以保证预报的精度。

由于计算MA(m)序列的预报值$\hat{x}_t(j)$($l\leqslant m$)需要对大量历史数据加权求和，而且每次获得新的数据后，必须与大量历史数据一道重新计算加权和，这就使预报过程中历史资料$x_1, x_2\cdots, x_K$的存贮量和计算量都随t而不断增大。这对于采样间隔短，资料积累又较丰富的连续实时预报，将会带来很大困难。因此下面要给出一种称为预报矢量法的实时递推方法，它可以达到减少存贮，缩短计算时间的目的。当然，对于那些非连续预报，或者虽然要求连续性预报，但采样间隔较长(如一小时或一天等)的实际问题就没有必要采取这种递推形式。

由式(6-1-13)及$G_i=b_i$($1\leqslant i\leqslant m$)，立即可得MA(m)序列的递推预报公式

$$\hat{X}_{t+1}^{(m)}=\begin{bmatrix}-b_1 & 1 & & 0\\ \vdots & & \ddots & \\ -b_{m-1} & 0 & & 1\\ -b_m & 0 & \cdots & 0\end{bmatrix}\hat{X}_t^{(m)}+\begin{bmatrix}b_1\\ b_2\\ \vdots\\ b_m\end{bmatrix}x_{t+1} \qquad (6\text{-}2\text{-}3)$$

式中$\widehat{X}_t^{(m)}$称做预报矢量，记为

$$\widehat{X}_{t_0}^{(m)}=[\hat{x}_t(1),\hat{x}_t(2),\cdots\hat{x}(m)]^T$$

式(6-2-3)连同$\hat{x}_t(l)=0, l>m$一起，描述了$MA(m)$序列以时刻t为起点，对未来进行的任意l步预报。式(6-2-3)中，对每一个时刻t，都只须存贮该时刻的预报矢量$\widehat{X}_t^{(m)}$（m个分量），递推初值$\widehat{X}_{t_0}^{(m)}$可通过格林函数方法或逆函数方法直接计算求出。当t_0距很远时，也可取$\widehat{X}_{t_0}^{(m)}=0$。可以证明，MA模型的可逆性保证了上述递推是渐近稳定的，即当t的值充分大以后，初值误差的影响可以忽略。因此如果进行长期递推预报，从初值开始经过充分长的一段时间后，预报值就可以用了。

下面举几个低阶AR，MA模型预报的例子，以说明纯自回归或滑动平均模型的预报。

例1．AR(1)的预报

模型为

$$x_t-\varphi_1 x_{t-1}=\varepsilon_t$$

由(6-2-1)可得

$$\hat{x}_t(l)=\varphi_1\hat{x}_t(l-1) \qquad l>0$$

初值为$\hat{x}_t(0)=x_t$，代入上式得到

$$\hat{x}_t(l)=\varphi_1^l x_t$$

例2．$MA(1)$的预报

模型为

$$x_t=\varepsilon_t+b_1\varepsilon_{t-1}$$

由(6-2-2)得知，MA(1)只有一步预报值是非零的，超过一步的预报值皆为零。由式(6-1-11)及MA序列的逆转形式知

$$I_j^{(1)}=I_j=-(-b_1)^j \qquad (j\geqslant 1)$$

因此得到

$$\hat{x}_t(1)=\sum_{j=1}^{\infty}I_j^{(1)}x_{t+1-j}=b_1\sum_{j=0}^{\infty}(-b_1)^j x_{t-j}\cong b_1\sum_{j=0}^{K}(-b_1)^j x_{t-j}$$

其中K的值由允许的误差确定，使得 $\sum_{i=K+1}^{\infty}|b_1^i x_{t-i}|$ 可以忽略不计。

如果用预报矢量递推法进行预报，由式(6-2-3)可得

$$\hat{x}_{t+1}(1)=-b_1\hat{x}_t(1)+b_1x_{t+1} \tag{6-2-4}$$

我们可以用直接算法(如前述逆函数方法)，求得某个$\hat{x}_{t_0}(1)$，然后用式(6-2-4)递推预报任意时刻的$x_t(1)$，$t>t_0$，当t增至较大时，$x_t(1)$将有很好的预报精度。

6.3. ARMA序列的预报方法

ARMA序列的预报与MA序列的预报类似，可以用逆函数方法进行预报，也可以用预报矢量递推，但计算比MA序列复杂。因为$MA(m)$序列超过m步的预报$\hat{x}_t(l)(l>m)$恒为零，而$ARMA$序列就无此特性。但由(6-1-8)可知，超过m步的预报值可以递推求出。因而对ARMA序列，我们也可以推导出预报矢量递推公式。

预报矢量仍定义为

$$\hat{X}_t^{(m)}=[\hat{x}_t(1),\hat{x}_t(2),\cdots,\hat{x}_t(m)]^T$$

由(6-1-8)，(6-1-13)可得

$$\hat{X}_{t+1}^{(m)}=\begin{pmatrix}-G_1 & 1 & & 0\\ -G_2 & & \ddots & \\ \vdots & & & \\ -G_{m-1} & 0 & & 1\\ -(G_m+a'_m) & -a'_{m-1} & \cdots & -a'_1\end{pmatrix}\hat{X}_t^{(m)}+\begin{pmatrix}G_1\\ G_2\\ \vdots\\ G_m\end{pmatrix}x_{t+1}+\begin{pmatrix}0\\ \vdots\\ 0\\ -\sum\limits_{j=m+1}^{n}a'_jx_{t+m-j+1}\end{pmatrix} \tag{6-3-1}$$

其中

$$a'_j=\begin{cases}a_j & 0<j\leqslant n\\ 0 & j>n\end{cases}$$

当$n\leqslant m$时式(6-3-1)中第三项为零。递推初值取作

$$\widehat{X}_{t_0}^{(m)}=[\hat{x}_{t_0}(1),\hat{x}_{t_0}(2),\cdots,\hat{x}_{t_0}(m)]^T$$

可由直接预报法计算得到，当然在可逆性条件满足时也可取$\widehat{X}_{t_0}^{(m)}=0$。

综合前述，就预报方式而言，ARMA序列的各种预报方法可分为直接预报和递推预报两大类。

递推预报即是利用式(6-3-1)预报矢量递推公式，从t步起不断向前延伸。但初始点t_0预报值$[\hat{x}_{t_0}(1),\hat{x}_{t_0}(2),\cdots,\hat{x}_{t_0}(m)]$仍由直接预报法求得。

直接预报是从模型本身出发进行预报。由于ARMA(n,m)模型具有通常的、传递的和逆转的三种等价形式，因此直接预报也可以分为由差分方程预报、由格林函数预报及由逆函数预报三类方法。现分别归纳叙述如下：

(1)由差分方程预报。设ARMA(n,m)模型写成

$$x_{t+l}+a_1x_{t+l-1}+\cdots+a_nx_{t+l-n}=\varepsilon_{t+l}+b_1\varepsilon_{t+l-1}+\cdots+b_m\varepsilon_{t+l-m}$$

上式两边在空间X_t上进行投影，注意$\mathscr{H}_t=\mathscr{E}_t$并利用式(6-1-6),(6-1-7)可得：

$$\hat{x}_t(l)+\sum_{j=1}^{l-1}a_j\hat{x}_t(l-j)=-\sum_{j=l}^{n}a_jx_{t-j+l}+\sum_{j=l}^{m}b_j\varepsilon_{t-j+l}\tag{6-3-2}$$

上式中若求和上限小于下限则认为其值为零。对于MA(m)序列而言，$a_j=0(j\geqslant1)$，若$l>m$则式(6-3-2)右边第二项为零，因此$\hat{x}_t(l)=0$(当$l>m$)，这与节6.2中的结论相符。对AR(n)序列而言，式(6-3-2)实际上化成式(6-2-1)。对一般ARMA序列而言，

当$l>m$时，式(6-3-2)也恰恰是式(6-1-8)。

式(6-3-2)是以$\hat{x}_t(1),\hat{x}_t(2),\cdots,\hat{x}_t(m),\cdots,\hat{x}_t(l)$为变元的线性方程组，它的系数矩阵是下三角阵（对角线上方元素皆为零），因此只需迭代求解。这在计算机上是很容易实现的，

(2)由格林函数预报。所谓用格林函数预报，就是按式(6-1-4)用白噪声和格林函数对未来值进行预报。格林函数按2.3节式(2-3-68)递推计算，白噪声ε_t则按下式迭代(初值取为零)

$$\varepsilon_t=x_t+a_{1\,t-1}+\cdots+a_n x_{t-n}-b_1\varepsilon_{t-1}-\cdots-b_m\varepsilon_{t-1}-\cdots-b_m\varepsilon_{t-m}$$

式(6-1-4)中包含无穷级数求和，实际计算时只能用充分多项的有穷和近似，即取足够大的T，使得$\sum_{j=T+1}^{\infty}|G_{j+l}\varepsilon_{t-j}|$按允许误差可忽略不计。

(3)由逆函数预报。所谓用逆函数进行预报，就是用现在和过去的观察数据以及逆函数，按式(6-1-10)对未来值作预报。逆函数按2.3节式(2-3-79)递推计算。

由式(6-1-10)可见，除AR序列外，用逆函数进行预报需要用到全部历史数据，因此实际只能实现近似的计算。

最后应当指出，本章所讨论的预报都是在平稳序列适合于某一$ARMA$模型的前提下进行的，对于较复杂的情况，如第八章将讨论的$ARIMA(n,d,m)$模型、季节模型、含确定性趋势的组合模型等，必需对原始观察序列采取适当的变换处理，使之平稳化，才能进行平稳预报。与此相应，如果在预报前对数据进行了平稳化处理，平稳预报后，一定要将预报值还原成与原序列相应的值，这样才能得出对原序列的预报值。

图6-3是对一个$ARMA(1,1)$过程，$x_t-0.8x_{t-1}=\varepsilon_t+0.4\varepsilon_{t-1}$，进行一步预报的例子，采用了逆函数和矢量递推两种方法(分别调用程序$INVERF$和$VECTF$)，过程的已知数据 $\{x_t, 1\leqslant$

$t \leqslant 150\}$ 和$t=101, \cdots, 150$处的一步预报值$\hat{x}_{100}(1), \hat{x}_{101}(1), \cdots, \hat{x}_{149}(1)$在表6-1和图6-3中均可看出，两种方法的预报结果十分接近。矢量递推法的预报初值取为$\hat{x}_{99}(1)=0.507$，它是由逆函数法直接计算得出的。矢量递推法和逆函数法的一步预报误差平方和(和真值相比)分别为0.99991和1.00929。

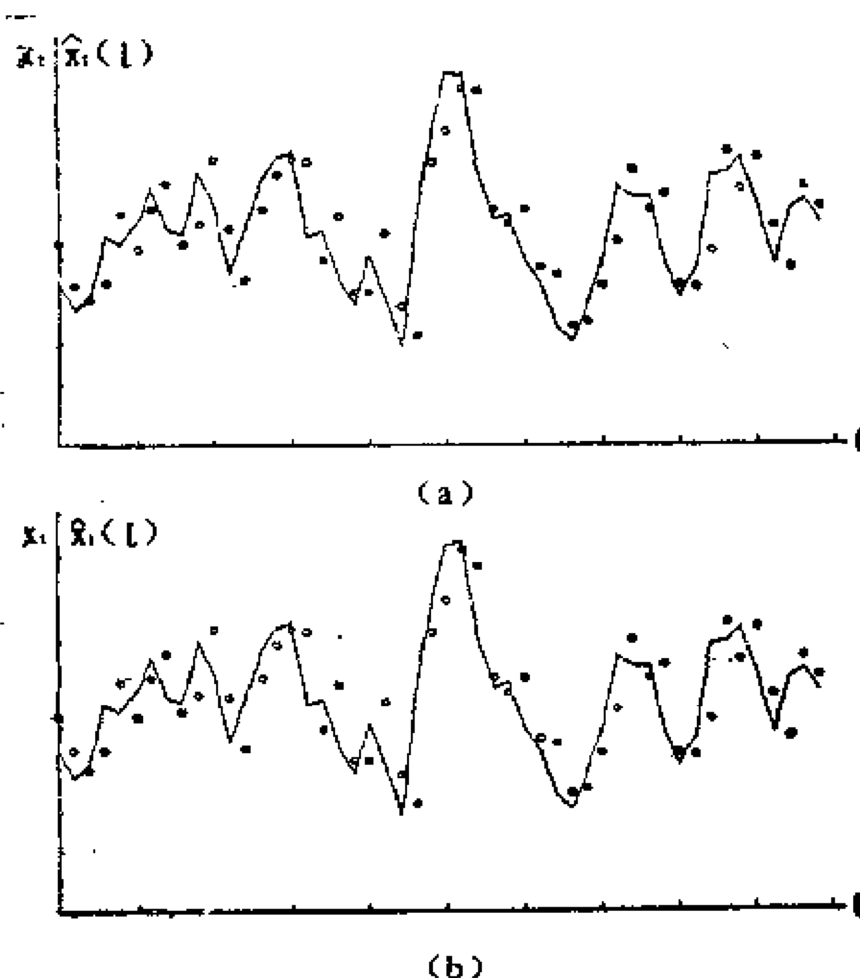


图6—3数据曲线和预报结果(以点表示)

a)逆函数法　　b)矢量递推法

表6-1　真值和预报值(1-真值，2-递函数法预报值，3-矢量递推法预报值)

1	-1.591	-2.228	-2.017	-.537	-.687	-.179	.588	-.345
2	-.669	-1.642	-2.017	-1.612	.001	-.826	.116	.659
3	-.816	-1.583	-2.041	-1.604	-.002	-.824	.115	.660
1	-.444	.962	.144	-1.379	-.425	.795	1.341	1.413
2	-.677	-.255	1.251	-.328	-1.523	.096	.917	1.256
3	-.678	-.262	1.260	-.331	-1.522	.099	.914	1.244
1	-.524	-.426	-1.551	-2.117	-.981	-1.957	-3.104	.028
2	1.193	-1.106	-.062	-1.835	-1.807	-.453	-2.171	-2.854
3	1.198	-1.108	-.068	-1.834	-1.807	-.454	-2.166	-2.858
1	2.008	3.249	3.342	1.034	-.134	.008	-1.061	-1.608
2	1.175	1.939	2.941	2.834	.107	-.204	.092	-1.277
3	1.177	1.939	3.124	2.761	.137	-.216	.096	-1.311
1	-2.664	-2.957	-2.192	-1.080	.633	.416	.414	-1.224
2	-1.419	-2.629	-2.497	-1.631	-.643	1.017	.091	.461
3	-1.405	-2.634	-2.495	-1.632	-.643	1.017	.093	.460
1	-1.963	-1.268	.894	.964	1.289	.213	-1.148	.116
2	-1.652	-1.693	-.841	1.406	.593	1.310	-.268	-1.273
3	-1.653	-1.695	-.843	1.411	.593	1.310	-.268	-2.270
1	.370	-.149						
2	.644	.184						
3	.647	.185						

6.4　时间序列的新息实时预报

前面几节讨论了$ARMA$序列的平稳预报，即是根据全部历史数据$x_t, x_{t-1}\cdots$，给出x_{t+l} $(l>0)$的线性最小方差预报$\hat{x}_t(l)$.然而在实际情况下，我们只能获得有限的资料．对于AR序列来说用有限资料就可以对未来值实现严格的平稳最小方差预报 。对MA或ARMA序列则只能计算出$\hat{x}_t(l)$的近似值，仅当t很大时，由近似所带来的误差才可能被忽略。在某些实际问题中，特别是那些要求连续进行实时预报的问题，常常希望根据实际上已获得的有限个

数据$x_t, x_{t-1}, \cdots, x_1$，给出对$x_{t+l}$严格的线性最小方差预报，并希望有一种存储量不随$t$的增加而增加的实时预报递推算式。本节将介绍的就是适应这种要求的新息预报方法。

1.新息预报的原理 与本章节6.1中所指的无穷序列x_t, x_{t-1}相对应，对于有限数据x_t，x_{t-1}，$\cdots$，x_1可平行地引入一系列的概念和记号。

设ARMA序列相应的模型是

$$x_t + a_1 x_{t-1} + \cdots + a_n x_{t-n} = \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_m \varepsilon_{t-m}$$

定义线性空间

$$\mathscr{X}_t^t = \{x : x = \sum_{j=1}^{t} c_j x_t, c_j \text{为实数}\}$$

为了与节6.1中平稳预报相区别，本节中用$\hat{x}_t^t(l)$表示用有限个历史数据$x_t, x_{t-1}, \cdots, x_1$对未来时刻的值$x_{t+l}$所做的线性最小方差预报。由定义自然有

$$\hat{x}_t^t(l) \in \mathscr{X}_t^t$$

且使方差$E[x_{t+l} - \hat{x}_t^t(l)]^2$达到最小。

记一步预报误差为

$$e_t = x_t - \hat{x}_{t-1}^t(1)$$

如果把$\hat{x}_{t-1}^t(1)$看成是t时刻以前的"老"数据$x_{t-1}, x_{t-2}, \cdots x_1$提供的关于$x_t$的最大信息，那么$e_t$就可以看作是从新来的数据$x_t$中得到的最新信息。因此我们称$e_1, e_2 \cdots$为$\{x_t\}$的新息序列。

定义

$$\mathscr{E}_t^t = \left\{ \xi : e = \sum_{j=1}^{t} d_j e_j, d_j \text{为实数} \right\}$$

与节6.1完全相似，应有

$$\mathscr{X}_t^t = \mathscr{E}_t^t$$

由新息预报的定义可知，$\hat{x}_2^t(l)$实质上就是x_{t+l}在线性空间X_t^t上

的正交投影。根据线性最小方差估计的性质容易得出

$$E[e_t x_j]=0, E[e_t e_j], \qquad j<t$$

这表明新息序列 $\{e_j, 1\leqslant j<t\}$ 是线性空间$\mathscr{H}_t^x$中的一组正交基。因此$\hat{x}_t^I(l)$可表示成$e_t, e_{t-1}, \cdots, e_1$的线性组合：

$$\hat{x}_t^I(l)=\sum_{j=1}^{t} J_{t+1,j} e_j \tag{6-4-1}$$

式中系数由极小化$E[x_{t+1}-\hat{x}_t^I(l)]^2$得到

$$J_{t+1,j}=(E[x_{t+1}e_j])(E[e_j^2])^{-1} \tag{6-4-2}$$

因此只要求出$E[x_{t+1}e_j]$与$E[e_j^2]$，便可计算新息预报值$\hat{x}_t^I(l)$。关于新息序列e_t的计算方法将在下节进一步介绍。

2.新息定理及实时预报的性质 新息预报方法基于以下定理

新息定理[13] 设$\{x_t\}$是平稳零均值序列，适合于如下ARMA(n,m)模型

$$x_t+a_1x_{t-1}+\cdots+a_nx_{t-n}=\varepsilon_t+b_1\varepsilon_{t-1}+\cdots+b_m\varepsilon_{t-m}$$

令M=max(n,m)，定义序列 $\{y_t\}$ 为

$$y_t=\begin{cases}x_t & t\leqslant M\\ x_t+a_1x_{t-1}+\cdots+a_nx_{t-n} & t>M\end{cases} \tag{6-4-3}$$

则新息序列$\{e_t^I\}$满足如下递推关系式

$$e_t=\begin{cases}y_1=x_1 & t=1\\ y_t-\sum\limits_{j=1}^{t-1}J_{tj}x_j & 1<t\leqslant M\\ y_t-\sum\limits_{j=t-m}^{t-1}J_{tj}e_j & t>M\end{cases} \tag{6-4-4}$$

其中

$$J_{tj}=\begin{cases}R_y(t,j)-\sum\limits_{i=1}^{j-1}J_{ji}J_{ti}R_e(i,i)][R_e(j,j)]^{-1} & 1\leqslant t\leqslant M,\ 1\leqslant j\leqslant t-1\\ [R_y(t,j)-\sum\limits_{i=j-m}^{j-1}J_{ji}J_{ti}R_e(i,i)][R_e(j,j)]^{-1} & t>M,\ t-m\leqslant j\leqslant t-1\end{cases} \tag{6-4-5}$$

式中

$$R_y(t,j)=E[y_t y_j],\quad R_e(j,j)=E[e_j^2]$$

且

$$R_{\tilde{e}}(t,t)=\begin{cases} R_y(1,1) & t=1 \\ R_y(t,t)-\sum\limits_{j=1}^{t-1} J_{t\,j}^2 R_e(j,j) & 1<t\leqslant M \\ R_y(t,t)-\sum\limits_{j=t-m}^{t-1} J_{t\,j}^2 R_e(j,j) & t>M \end{cases}$$

上列各式中，当求和上限小于下限时，约定其和值为零。下面证明这个定理。

由式(6-4-3)可得序列$\{y_t\}$，我们定义线性空间

$$\mathscr{Y}'_t=\{y:y=\sum_{j=1}^{t}\mu_j y_j,\ \mu_j \text{为实数}\}$$

由定义容易看出

$$\mathscr{Y}'_t=\mathscr{X}'_t=\mathscr{E}'_t$$

当$t\leqslant M$时

$$y_t-\hat{y}'_{t-1}(1)=x_t-\hat{x}'_{t-1}(1)=e_t$$

当$t>M$时，利用$\hat{y}'_{t-1}(1)=\hat{E}[y_t|\mathscr{X}'_{t-1}]$

$$\begin{aligned} y_t-\hat{y}'_{t-1}(1) &=\sum_{j=0}^{n}a_j x_{t-j}-\hat{E}[\sum_{j=0}^{n}a_t x_{t-j}|\mathscr{X}'_{t-1}] \\ &=\sum_{j=0}^{n}a_j x_{t-j}-(\hat{x}'_{t-1}(1)+\sum_{j=1}^{n}a_j x_{t-j}) \\ &=x_t-\hat{x}'_t(1)=e_t \end{aligned}$$

由此可见$\{y_t\}$与$\{x_t\}$有相同的新息序列$\{e_t\}$。类似于式(6-4-1)和(6-4-2)也可以得到对y_t的新息预报：

$$\hat{y}'_t(l)=\sum_{j=1}^{t}J_{t+l;j}e_j \tag{6-4-6}$$

$$J_{t+1,j}=(E[y_{t+1}e_j])(Ee_j^2)^{-1} \tag{6-4-7}$$

由 $\{y_t\}$ 的定义知，当$t>M$时

$$y_t=\sum_{j=0}^{n}a_j x_{t-j}=\sum_{j=0}^{m}b_j\varepsilon_{t-j}$$

即当$t>M$时，$\{y_t\}$ 是一个m阶滑动平均序列，因此有

$$E[y_t y_j]=0 \qquad t>M,\ 1\leqslant j<t-m$$

这说明y_t与$\mathscr{Y}'_{t-m-1}=\mathscr{E}'_{t-m-1}$正交，所以当$t>M$时，

$$E[y_t e_j]=0 \qquad 1\leqslant j<t-m$$

亦即当$t>M$时y_t与$e_1,\cdots e_{t-m-1}$不相关，因而事实上，用$e_1, e_2, \cdots e_{t-1}$的线性组合表示$\hat{y}'_{t-1}(1)$应有

$$\hat{y}'_{t-1}(1)=\begin{cases}\sum\limits_{j=1}^{t-1}J_{tj}e_j & 当t\leqslant M\\ \sum\limits_{j=\max(t-m,1)}^{t-1}J_{tj}e_j & 当t>M\end{cases} \tag{6-4-8}$$

考虑到$\hat{y}'_{t-1}(1)=y_t-e_t$，式(6-4-8)显然就是式(6-4-4)，为证明(6-4-5)式，先求$E[e_j y_t]$，利用$y_t=e_t+\hat{y}'_{t-1}(1)$及式(6-4-8)可得，当$t\leqslant M$时

$$E[e_j y_t]=E[e_j(e_t+\sum_{i=1}^{t-1}J_{ti}e_i)]=\begin{cases}J_{tj}R_e(j,j) & 1\leqslant j\leqslant t-1\\ R_e(j,j) & j=t\\ 0 & j>t\end{cases}$$

当$t>m$时

$$E[e_j y_t]=E[e_j(e_t+\sum_{i=\max(t-m,1)}^{t-1}J_{ti}e_i)]$$

$$=\begin{cases}J_{tj}R_e(j,j) & t-m\leqslant j<t\\ R_e(j,j) & j=t\\ 0 & j>t或1\leqslant j\leqslant t-m\end{cases}$$

根据以上二式可知，当$1\leqslant j\leqslant t\leqslant M$时

$$J_{tj}R_e(j,j)=E[e_jy_t]=E[(y_j-\sum_{i=1}^{j-1}J_{ji}e_i)y_t]$$

$$=R_y(t,j)-\sum_{i=1}^{j-1}J_{ji}J_{ti}R_e(i,i)$$

当$t>m$且$t-m\leqslant j<t-1$时

$$J_{tj}R_e(j,j)=E[e_jy_t]=E[(y_j-\sum_{i=\max(j-m,1)}^{j-1}J_{ji}e_i)y_t]$$

$$=R_y(t,j)-\sum_{i=\max(j-m,1)}^{j-1}J_{ji}J_{ti}R_e(i,i)$$

上式两端乘以$[R_e(j,j)]^{-1}$便得到式(6-4-5)。定理证毕。

由式(6-4-4)，和式(6-4-3)可得

$$e_{t+l}=\begin{cases}x_{t+l}-\sum\limits_{j=1}^{t+l-1}J_{t+l,j}e_j & t+l\leqslant M\\ x_{t+l}+\sum\limits_{j=1}^{n}a_jx_{t+l-j}-\sum\limits_{j=t+l-m}^{t+l-1}J_{t+l,j}e_j & t+l>M\end{cases}$$

上式经调整，并在线性空间X_t^l上投影后，便可得到l步实时预报公式

$$\hat{x}_t'(l)=\begin{cases}\sum\limits_{j=1}^{t}J_{t+l,j}e_j & t+l\leqslant M\\ -\sum\limits_{j=1}^{n}a_j\hat{x}_t'(l-j)+\sum\limits_{j=t+l-m}^{t}J_{t+l,j}e_j & t+l>M\end{cases}\qquad(6\text{-}4\text{-}9)$$

式(6-4-9)中，当$l-j\leqslant 0$时取$\hat{x}_t'(l-j)=x_{t+l-j}$

由新息定理可以看出，新息序列$\{e_t\}$是随着观察数据x_t输入，经递推得到的。在进行新息预报过程中，需要存储的量是$\hat{x}_t'(l-j)$（$t+l>M$时，$1\leqslant j\leqslant n$），$J_{t+l,j}$，e_j（当$t+l\leqslant M$时，$j=1,2\cdots,t+l-1$当$t+l>M$时，$j=t+l-m,\cdots,t+l-1$），显然

存储量并不随t而增大。另外$R_{\gamma}(t,j)$可由样本相关$\hat{R}_x(j)=E[x_t x_{t-j}]$及模型自回归参数$a_1,\cdots,a_n$直接算出，而且还可指出，$R_{\gamma}(t,j)$只是$t-j$的函数(注意$1\leqslant t-j\leqslant m$)，因此求$R_{\gamma}(t,j)$的计算量是不大的。总之，新息预报虽然公式较复杂，但占用的内存是有限的，并不随t而增长，而且每步实时预报是用递推计算，在计算机上实现很方便。特别是MA序列，由新息预报公式可以看出，只要能判断出MA模型的阶数，不必计算出滑动平均参数就可以递推进行新息预报，因此对于MA模型新息预报是特别方便的。

在式(6-4-9)中，当$l>m$时，右边第二个和式为零，该式成为

$$\sum_{j=1}^{n} a_j \hat{x}'_t(l-j)=0 \qquad l>m$$

对于$AR(n)$序列，$m=0$，$M=\max(n,m)=n$,即有

$$\sum_{j=0}^{n} a_j \hat{x}'_j(l-j)=0, \qquad l>0, \ t>n$$

递推初值为$\hat{x}'_t(j)=x_{t+j}(j=0,-1,\cdots,-n+1)$。将此式与节6.2中AR序列的平稳线性最小方差预报对照，可以看出当$t>n$时，AR序列的新息预报与平稳预报与平稳预报是完全相同的。因此

$$e_t=x_t-\hat{x}_{t-1}(1)\equiv x_t-\hat{x}_{t-1}(1)=\varepsilon_t, \ t>n$$

对MA或ARMA序列，一般$\varepsilon_t\neq e_t$，但不难证明，平稳预报与新息预报有渐近相等的一步预报误差方差，即：

$$E(e_t^2)\to\sigma_\varepsilon^2 \quad (t\to\infty)$$

而且一步预报误差本身也是渐近相等的：

$$E(e_t-\varepsilon_t)^2\to 0 \qquad (t\to\infty)$$

上述分析表明，无论是AR,MA序列，还是ARMA序列，当t很大时，实时预报与平稳预报都渐近趋于一致。因此在实际应用中遇到连续预报问题时，若要求从较少的数据开始预报，并希望尽

可能给出好的预报值，那么开始一段，可以使用实时预报方法。当t增大到一定数值后，为了减少每步的计算量，改用平稳预报的递推方法进行预报。这样把两种方法结合起来使用，既可以提高预报精度，又可以节省计算量。

第七章　多变量时间序列

如果我们不仅是观察单个过程x_t，而是同时观察几个过程——$x_{1,t}$，$x_{2,t}$，…，$x_{p,t}$，或者说当t为标量而x为向量时，需要分析多变量时间序列$\mathbf{x}_t=(x_{1,t}\ x_{2,t}\cdots x_{p,t})^T$。例如在工程中要研究电流和电压同时随时间变化的情况，在化学变化中要分析压力、温度和体积的变化关系，在气象预报分析时，需要同时考虑该地区的雨量、气温和气压等纪录资料。在研究较复杂的控制系统时，也会遇到多输入和多输出的关系问题。当然，可以把它们分别看作是单变量过程来研究各个量本身，尽管这也能提供各个量所包含的信息，但却不能揭示这些量之间可能存在着的相互关系。正如在概率论中分析几个随机变量的关系时，不仅要知道边缘分布，而且要知道联合分布一样，我们需要有描述多变量过程的一些二阶联合特性，这就是本章所要介绍的互协方差（或互相关）函数、互谱以及相应的多变量ARMA模型等。

7.1　多变量平稳过程的相关和谱特性

1.双变量过程的相关特性　为引出一般的多变量的情况，先看双变量随机过程$\{x_{1,t}\}$和$\{x_{2,t}\}$，$t=0,\pm1,\pm2,\cdots$。如果

（1）$\{x_{1,t}\}$和$\{x_{2,t}\}$作为各自的单变量过程来说是平稳（见第二章定义），而且

（2）$\mathrm{Cov}(x_{1,t},x_{2,s})$仅是$(s-t)$的函数。

则称$\{x_{1,t};\ x_{2,t}\}$是平稳双变量过程，或者说$\{x_{1,t}\}$,$\{x_{2,t}\}$是联合平稳的。这时可以用通常的办法定义$\{x_{1,t}\}$,$\{x_{2,t}\}$的自协方差函数：

$$R_{11}(r)=E((x_{1,t}-\mu_1)^*(x_{1,t+r}-\mu_1))$$
$$R_{22}(r)=E((x_{2,t}-\mu_2)^*(x_{2,t+r}-\mu_2)) \qquad (7\text{-}1\text{-}1)$$
其中$\mu_1=E[x_{1,t}]$，$\mu_2=E[x_{2,t}]$，以上定义不论$x_{1,t}$，$x_{2,t}$为实过程或复过程都适用。相应的自相关函数为
$$\rho_{11}(r)=R_{11}(r)/R_{11}(o)$$
$$\rho_{22}(r)=R_{22}(r)/R_{22}(o) \qquad (7\text{-}1\text{-}2)$$
自相关函数只描述各个过程内在的相关结构，而不同过程之间的相关结构则须用下面定义的互协方差(互相关)函数描述。$\{x_{1,t}\}$ $\{x_{2,t}\}$的互协方差函数为
$$R_{21}(r)=\text{Cov}[x_{1,t},x_{2,t+r}]=E((x_{1,t}-\mu_1)^*(x_{2,t+r}-\mu_2)) \qquad (7\text{-}1\text{-}3)$$
根据前面提到的条件（2），$R_{21}(r)$只和r有关而与t无关。

互相关函数定义为
$$\rho_{21}(r)=R_{21}(r)/(R_{11}(o)R_{22}(o))^{1/2} \qquad (7\text{-}1\text{-}4)$$
$R_{21}(r)$表示的是"$x_{1,t}$先于$x_{2,t}$"的互协方差函数。对称地，可以有"$x_{2,t}$先于$x_{1,t}$"的互协方差函数$R_{12}(r)$及相应的互相关函数$\rho_{12}(r)$。它们定义为
$$R_{12}(r)=E((x_{2,t}-\mu_2)^*(x_{1,t+r}-\mu_1)) \qquad (7\text{-}1\text{-}5)$$
$$\rho_{12}(r)=R_{12}(r)/(R_{11}(o)R_{22}(o))^{1/2} \qquad (7\text{-}1\text{-}6)$$
由于
$$R_{12}(r)=R_{21}^*(-r) \qquad (7\text{-}1\text{-}7)$$
因此R_{12}和R_{21}所包含的两个过程相互关系的信息是相同的。

双变量过程的全部互协方差特性可用矩阵
$$R(r)=\begin{pmatrix} R_{11}(r) & R_{12}(r) \\ R_{21}(r) & R_{22}(r) \end{pmatrix} \qquad (7\text{-}1\text{-}8)$$
归纳起来。称$R(r)$为迟后r的协方差阵，类似方法可以定义迟后r的互相关阵$\rho(r)$。显然有

$$|\rho_{12}(r)| \leqslant 1 \quad 任何r。$$

对于实过程，$\rho_{11}(r)$ 和$\rho_{22}(r)$ 都是偶函数且$r=0$处为1。但$\rho_{12}(r)$ 一般不是偶函数而且最大值可在任意r处。图7-1 表示两个实过程样本的互相关函数。

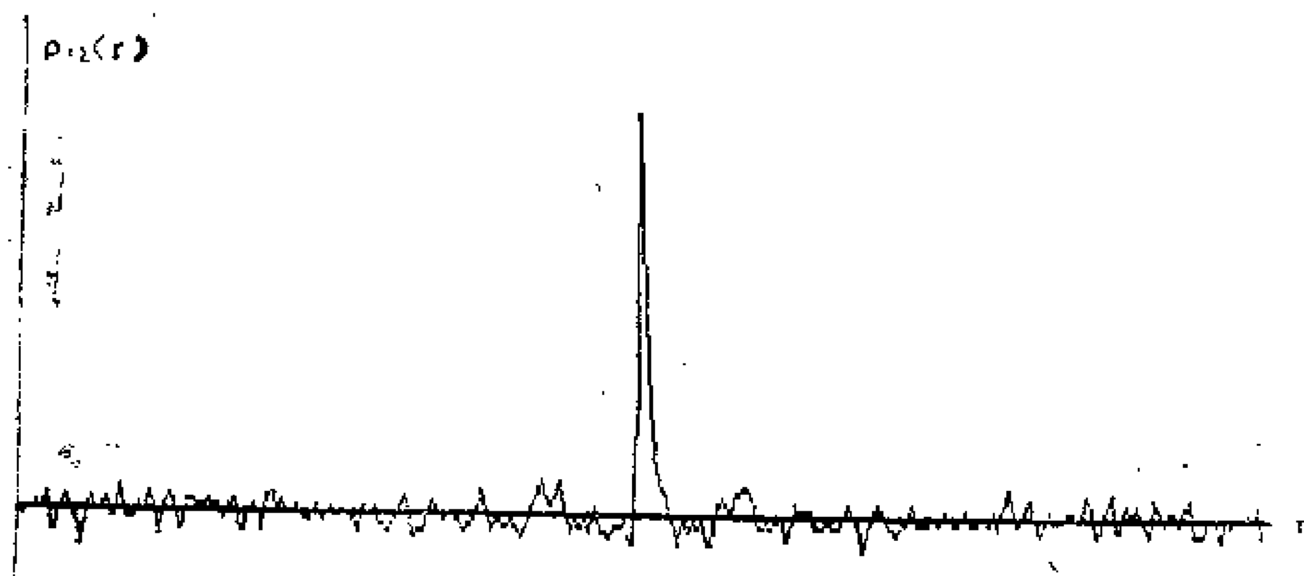


图7-1　一个互相关函数样本

由于$x_{1,t}$和$x_{2,t}$的均值都是常数，可令$\mu_1=\mu_2=0$而不失一般性。以后假定$x_{1,t}$和$x_{2,t}$为零均值过程进行讨论。

2. 双变量过程的谱特性　如果$\{x_{1,t}\}\{x_{2,t}\}$均具有纯粹连续的谱密度函数—$G_{11}(f)$和$G_{22}(f)$。它们是各自自协方差函数的富氏变换

$$G_{11}(f)=\sum_{r=-\infty}^{\infty} R_{11}(r)e^{-j2\pi fr} \tag{7-1-9}$$

$$G_{22}(f)=\sum_{r=-\infty}^{\infty} R_{22}(r)e^{-j2\pi fr} \tag{7-1-10}$$

互协方差函数$R_{21}(r)$的富氏变换

$$G_{21}(f)=\sum_{\tau=-\infty}^{\infty} R_{21}(\tau)e^{-j2\pi f\tau} \tag{7-1-11}$$

称为互谱密度函数或简称“互谱”。如果$R_{21}(\tau)$是绝对可加的，即 $\sum_{\tau=-\infty}^{\infty}|R_{21}(\tau)|<\infty$，则$G_{21}(f)$在任何$f$下存在。

利用随机过程的谱展式可以对$G_{21}(f)$作出物理解释。由2-4-5节知，因$\{x_{1,t}\}$，$\{x_{2,t}\}$均为平稳，可以表示成

$$x_{1,t}=\int_{-1/2}^{1/2}e^{j2\pi ft}dZ_1(f) \tag{7-1-12}$$

$$x_{2,t}=\int_{-1/2}^{1/2}e^{j2\pi ft}dZ_2(f) \tag{7-1-13}$$

其中$\{dZ_1(f)\}$和$\{dZ_2(f)\}$均为正交过程，即当$f\neq f'$时
$E[dZ_1(f)dZ_1(f')]=E[dZ_2(f)dZ_2(f')]=0$，将式(7-1-12)(7-1-13)代入式(7-1-3)，由于$\mu_1=\mu_2=0$，可得

$$R_{21}(\tau)=\int_{-1/2}^{1/2}\int_{-1/2}^{1/2}e^{-j2\pi ft}e^{j2\pi f'(t+\tau)}E[dZ_1^*(f)dZ_2(f')] \tag{7-1-14}$$

因$R_{21}(\tau)$只是τ的函数，故上式右边必与t无关，这显然只有当

$$E[dZ_1^*(f)dZ_2(f')]=0 \quad f\neq f'$$

时才成立，亦即$\{dZ_1(f)\}$和$\{dZ_2(f)\}$不仅各自正交而且互为正交。根据这一性质，式(7-1-14)可简化为

$$R_{21}(\tau)=\int_{-1/2}^{1/2}e^{j2\pi f\tau}E[dZ_1^*(f)dZ_2(f)] \tag{7-1-15}$$

类似地

$$R_{11}(\tau)=\int_{-1/2}^{1/2}e^{j2\pi f\tau}E[|dZ_1(f)|^2] \tag{7-1-16}$$

$$R_{22}(r)=\int_{-1/2}^{1/2}e^{j2\pi fr}E\{|dZ_2(f)|^2\} \qquad (7\text{-}1\text{-}17)$$

由式（7-1-9）（7-1-10）（7-1-11）取反变换得

$$R_{21}(r)=\int_{-1/2}^{1/2}e^{j2\pi fr}G_{21}(f)df \qquad (7\text{-}1\text{-}18)$$

$$R_{11}(r)=\int_{-1/2}^{1/2}e^{j2\pi fr}G_{11}(f)df \qquad (7\text{-}1\text{-}19)$$

$$R_{22}(r)=\int_{-1/2}^{1/2}e^{j2\pi fr}G_{22}(f)df \qquad (7\text{-}1\text{-}20)$$

将这三式同式（7-1-15）（7-1-16）（7-1-17）比较可见

$$G_{11}(f)df=E\{|dZ_1(f)|^2\} \qquad (7\text{-}1\text{-}21)$$

$$G_{22}(f)df=E\{|dZ_2(f)|^2\} \qquad (7\text{-}1\text{-}22)$$

$$G_{21}(f)df=E\{dZ_1^*(f)dZ_2(f)\} \qquad (7\text{-}1\text{-}23)$$

可见，（非归一化）互谱密度函数可以这样解释：$\{G_{21}(f)df\}$表示$x_{1,t}$和$x_{2,t}$中$e^{j2\pi ft}$系数乘积的平均值。在单变量过程中$\{G_{11}(f)df\}$表示$x_{1,t}$中$e^{j2\pi ft}$系数平方的平均值，或者频率为f的分量的功率。我们也可以说，$G_{11}(f)df$和$G_{22}(f)df$分别表示$dZ_1(f)$和$dZ_2(f)$间的协方差。

有时为了区别不同的谱函数，称$G_{11}(f)$，$G_{22}(f)$为自谱密度函数，或简称“自谱”。

对$R_{12}(r)$取富氏变换得

$$G_{12}(f)=\sum_{r=-\infty}^{\infty}R_{12}(r)e^{-j2\pi fr} \qquad (7\text{-}1\text{-}24)$$

将双变量过程的谱特性归纳成如下谱阵

$$G(f)=\begin{pmatrix}G_{11}(f) & G_{12}(f)\\ G_{21}(f) & G_{22}(f)\end{pmatrix} \qquad (7\text{-}1\text{-}25)$$

相应于归一化的自谱$S_{11}(f)$，$S_{22}(f)$（它们是$\rho_{11}(r)$和$\rho_{22}(r)$

的富氏变换），可定义归一化的互谱为

$$S_{12}(f)=\sum_{r=-\infty}^{\infty}\rho_{12}(r)e^{-j2\pi fr} \tag{7-1-26}$$

类似可定义$S_{21}(f)$

由于$R_{12}(r)$不一定是偶函数，因此互谱密度函数$G_{12}(f)$一般为复数，即可表为

$$G_{12}(f)=P_{12}(f)-jQ_{12}(f) \tag{7-1-27}$$

其中$P_{12}(f)$和$-Q_{12}(f)$分别为$G_{12}(f)$的实部和虚部。称$P_{12}(f)$为$x_{1,t}$和$x_{2,t}$的"协谱"（co-*spectrum*），$Q_{12}(f)$为$x_{1,t}$和$x_{2,t}$的"转象谱"（*quadrature spectrum*）。

对实过程而言，$R_{12}(r)$可分解为偶函数部分—$[R_{12}(r)+R_{12}(-r)]/2$和奇函数部分$[R_{12}(r)-R_{12}(-r)]/2$，故

$$P_{12}(f)=\sum_{r=-\infty}^{\infty}\frac{1}{2}[R_{12}(r)+R_{12}(-r)]\cos 2\pi fr \tag{7-1-28}$$

$$Q_{12}(f)=\sum_{r=-\infty}^{\infty}\frac{1}{2}[R_{12}(r)-R_{12}(-r)]\sin 2\pi fr \tag{7-1-29}$$

这里$P_{12}(f)$是f的偶函数，$Q_{12}(f)$是f的奇函数。

另一种表示$G_{12}(f)$的方式是写成"极坐标"形式即

$$G_{12}(f)=M_{12}(f)e^{j\psi_{12}(f)} \tag{7-1-30}$$

其中

$$M_{12}(f)=|G_{12}(f)|=[P_{12}^2(f)+Q_{12}^2(f)]^{1/2} \tag{7-1-31}$$

$$\psi_{12}(f)=\tan^{-1}[-Q_{12}(f)/P_{12}(f)] \tag{7-1-32}$$

$M_{12}(f)$称为"互幅谱"，$\phi_{12}(f)$称为"互相谱"。

还有一个常用的谱特性函数是在频率f处的"复相干"（Complex Coherency），它定义为

$$W_{12}(f)=\frac{G_{12}(f)}{[G_{11}(f)G_{22}(f)]^{1/2}} \tag{7-1-33}$$

它的模$|W_{12}(f)|$定义为"相干"，$|W_{12}(f)|$曲线称为"相干谱"。（有的文献中称$|W_{12}(f)|^2$为相干或平方相干）

利用式（7-1-21）～（7-1-23）可将$W_{12}(f)$写成

$$W_{12}(f)=\frac{\mathrm{Cov}[dZ_1(f),dZ_2(f)]}{\{\mathrm{Var}[dZ_1(f)]\mathrm{Var}[dZ_2(f)]\}^{1/2}}$$

$$\tag{7-1-34}$$

因此$W_{12}(f)$可以认为是$x_{1,t}$和$x_{2,t}$中频率为f的分量的随机系数之间的相关系数。它显然是无量纲的，且

$$0 \leqslant |W_{12}(f)| \leqslant 1 \tag{7-1-35}$$

由此可知，$|W_{12}(f)|$和1的接近程度标志着随机系数（在频率f处）线性相关的程度，而$|W_{12}(f)|$在全部f上的表现形式决定了$x_{1,t}$和$x_{2,t}$二过程的线性相关程度。因此简单地说，$W_{12}(f)$是频率域中的相关系数。

我们知道，当对两个随机变量作线性变换时，它们之间的相关系数不变。同样，对$x_{1,t}$和$x_{2,t}$作线性变换也不改变相干。如变换为

$$x'_{1,t}=\Phi_1(B)x_{1,t}=\sum_{i=-\infty}^{\infty}\phi_{1,i}\,x_{1,t-i}$$

$$x'_{2,t}=\Phi_2(B)x_{2,t}=\sum_{i=-\infty}^{\infty}\phi_{2,i}\,x_{2,t-i}$$

令

$$\Phi(f)\equiv\Phi(B)\,|_{B=\exp(-i2\pi f)}$$

则有

$$dZ'_1(f)=\Phi_1(f)dZ_1(f)$$

$$dZ'_2(f)=\Phi_2(f)dZ_2(f)$$

故

$$G'_{11}(f)=|\Phi_1(f)|^2G_{11}(f)$$

$$G'_{22}(f)=|\Phi_2(f)|^2G_{22}(f)$$

$$G'_{12}(f)=\Phi_1(f)\Phi_2(f)G_{12}(f)$$

由式（7-1-33）可得

$$W'_{12}(f)=W_{12}(f)$$

必须指出，上述结果的导出是基于这样的假定：$x_{1,t}$和$x_{2,t}$都是平稳过程，它的谱和互谱在所有f下均存在。如果过程带有“趋势”或“季节性”分量，则上述结果不适用。

3.双变量过程举例

（1）不相关过程。这时对所有的r和t有$\mathrm{Cov}(x_{1,t},x_{2,r})=0$，即

$$R_{12}(r)=0 \qquad 所有r$$

因此 $\quad G_{12}(f)=0 \qquad$ 所有f

相应的所有互谱特性函数均为零，即任何f下$P_{12}(f)=0$，$Q_{12}(f)=0$，$M_{12}(f)=0$，$W_{12}(f)=0$。

（2）线性回归。若$x_{1,t}$和$x_{2,t}$满足线性回归关系

$$x_{1,t}=ax_{2,t}+\varepsilon_t$$

其中ε_t为纯随机过程且与$x_{2,t}$无关，故

$$R_{12}(\tau)=aR_{22}(\tau)$$

于是 $$G_{12}(f)=aG_{22}(f)$$

$$G_{11}(f)=a^2G_{22}(f)+G_\varepsilon(f)$$

因此该模型有

$$P_{12}(f)=aG_{22}(f),\quad Q_{12}(f)=0,$$

$$M_{12}(f)=aG_{22}(f),\quad \psi_{12}(f)=0$$

$$W_{12}(f)=\frac{1}{\sqrt{1+G_\varepsilon(f)/[a^2G_{22}(f)]}}$$

可见，若$\varepsilon_t\equiv 0$，则在任何f下$W_{12}(f)\equiv 1$，因为没有ε_t时$x_{1,t}$和$x_{2,t}$为准确的线性关系，故$x_{1,t}$和$x_{2,t}$的相干总是1。

（3）带滞后的线性回归。$x_{1,t}$和$x_{2,t}$满足线性回归关系，但有d时间单位的滞后，即

$$x_{1,t}=ax_{2,t-d}+\varepsilon_t$$

这时$x_{1,t}$和$x_{2,t}$的互相关是自相关的位移（例见图7-2）

$$R_{12}(\tau)=aR_{22}(\tau-d)$$

$$G_{12}(f)=a\sum_{\tau=-\infty}^{\infty}R_{22}(\tau-d)e^{-j2\pi f\tau}$$

$$=ae^{-j2\pi fd}G_{22}(f)$$

$$G_{11}(f)=a^2G_{22}(f)+G_\varepsilon(f)$$

$$P_{12}(f)=(a\cos 2\pi fd)G_{22}(f)$$

$$Q_{12}(f)=(a\sin 2\pi fd)G_{22}(f)$$

$$M_{12}(f)=aG_{22}(f)$$

$$\psi_{12}(f)=-2\pi fd$$

$W_{12}(f)$和上例相同

这里特别要指出的是相谱$\psi_{12}(f)$，在有时间滞后的情况下，

相谱是频率的线性函数，其斜率代表滞后的大小。一般情况下，对任何双变量平稳过程常用“包络滞后”，它定义为

$$D(f)=-d\psi_{12}(f)/df$$

以上推导$G_{12}(f)$是直接利用式（7-1-24），实际上利用式（7-1-23）可以更简捷地得出$G_{12}(f)$。因为由

$$x_{2,t-d}=\int_{-1/2}^{1/2} e^{j2\pi f(t-d)}dZ_2(f)$$

$$=\int_{-1/2}^{1/2} e^{j2\pi ft}[e^{-j2\pi fd}dZ_2(f)]$$

可得

$$dZ_1(f)=ae^{-j2\pi fd}dZ_2(f)+dZ_\varepsilon(f)$$

这里考虑到$x_{2,t}$和ε_t是无关的，因此$dZ_2(f)$和$dZ_\varepsilon(f)$也不相关，由式（7-1-23）

$$G_{12}(f)=E[dZ_1(f)dZ_2^*(f)]/df$$

$$=ae^{-j2\pi fd}E[|dZ_2(f)|^2]/df$$

$$=ae^{-j2\pi fd}G_{22}(f)$$



图7-2　两个滞后时间序列的互相关

（4）简单反馈系统（图7-3）

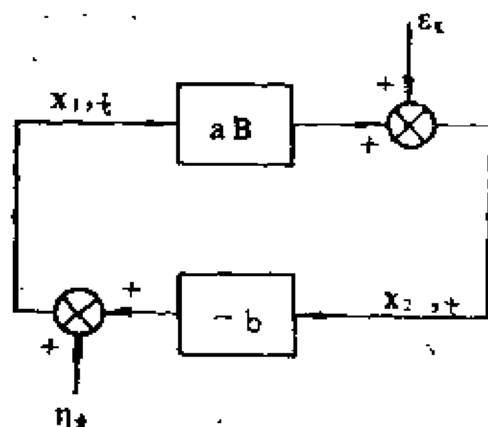


图7-3 简单反馈系统图

$$x_{2,t}=ax_{1,t-1}+\varepsilon_t \qquad x_{1,t}=-bx_{2,t}+\eta_t$$

其中$a>0$,$b>0$,ε_t和η_t是两个无关的白噪声过程。利用上例中提到的方法可以列出

$$dZ_2(f)=ae^{-j2\pi f}dZ_1(f)+dZ_\varepsilon(f)$$
$$dZ_1(f)=-bdZ_2(f)+dZ_\eta(f)$$

解这两个方程得

$$dZ_1(f)=\frac{dZ_\eta(f)-bdZ_\varepsilon(f)}{1+abe^{-j2\pi f}}$$

$$dZ_2(f)=\frac{ae^{-j2\pi f}dZ_\eta(f)+dZ_\varepsilon(f)}{1+abe^{-j2\pi f}}$$

将它们代入式(7-1-21)～(7-1-23)，令$\sigma_\eta^2=\text{Var}(\eta_t)$,$\sigma_\varepsilon^2=\text{Var}(\varepsilon_t)$，并考虑到$dZ_\eta(f)$和$dZ_\varepsilon(f)$不相关，得

$$G_{11}(f)=\frac{\sigma_\eta^2+b^2\sigma_\varepsilon^2}{|1+abe^{-j2\pi f}|^2}$$

$$=\frac{\sigma_\eta^2+b^2\sigma_\varepsilon^2}{1+a^2b^2+2ab\cos 2\pi f}$$

$$G_{22}(f)=\frac{a^2\sigma_\eta^2+\sigma_\varepsilon^2}{|1+abe^{-j2\pi f}|^2}$$

$$=\frac{a^2\sigma_\eta^2+\sigma_\varepsilon^2}{1+a^2b^2+2ab\cos 2\pi f}$$

$$G_{12}(f)=\frac{ae^{-j2\pi f}\sigma_\eta^2-b\sigma_\varepsilon^2}{|1+abe^{-j2\pi f}|^2}$$

$$=\frac{ae^{-j2\pi f}\sigma_\eta^2-b\sigma_\varepsilon^2}{1+a^2b^2+2ab\cos 2\pi f}$$

还可求得

$$P_{12}(f)=\frac{a\sigma_\eta^2\cos 2\pi f-b\sigma_\varepsilon^2}{1+a^2b^2+2ab\cos 2\pi f}$$

$$Q_{12}(f)=\frac{a\sigma_\eta^2\sin 2\pi f}{1+a^2b^2+2ab\cos 2\pi f}$$

$$M_{12}(f)=\frac{[(a\sigma_\eta^2\cos 2\pi f-b\sigma_\varepsilon^2)^2+a^2\sigma_\eta^4\sin^2 2\pi f]^{1/2}}{1+a^2b^2+2ab\cos 2\pi f}$$

$$\psi_{12}(f)=tan^{-1}\left[\frac{a\sigma_\eta^2\sin 2\pi f}{a\sigma_\eta^2\cos 2\pi f-b\sigma_\varepsilon^2}\right]$$

附带指出，上述反馈系统的实际背景是一个商品的价格和供应关系的简单经济模型。若以$x_{1,t}$和$x_{2,t}$分别表示某种商品在t时刻的价格和供应量，它可以用如下差分方程表示

$$x_{1,t}=\alpha-\beta x_{2,t}+\eta_t$$
$$x_{2,t}=\gamma+\delta x_{1,t-1}+\varepsilon_t$$

（其中$\alpha,\beta,\gamma,\delta$为实常数），如果将它中心化（均值为零）就可得到前面提到的模型。

4. 一般多变量过程 双变量过程的概念不难推广到多变量

的情况。若有p个零均值离散过程$x_{1,t}, x_{2,t}, \cdots, x_{p,t}$，则定义迟后$r$的协方差阵为

$$R(r)=[R_{ij}(r)] \qquad i,j=1,\cdots,p \tag{7-1-36}$$

其中 $$R_{ij}(r)=E[x_{j,t}^{*}x_{i,t+r}] \tag{7-1-37}$$

当$i=j$时$R_{ii}(r)$表示$x_{i,t}$的自协方差函数，$i\neq j$时$R_{ij}(r)$表示$x_{i,t}$和$x_{j,t}$之间的互协方差函数（$x_{j,t}$先于$x_{i,t}$）。假若p个过程是联合平稳的，即所有的i和j下，$R_{ij}(r)$仅是r的函数（与t无关），则谱展式为

$$x_{i,t}=\int_{-1/2}^{1/2}e^{j2\pi ft}dZ_i(f) \qquad i=1,\cdots,p \tag{7-1-38}$$

其中$\{dZ_i(f)\}$是正交和互正交的。将式（7-1-38）代入式（7-1-37）可得协方差函数的谱展式

$$R_{ij}(r)=\int_{-1/2}^{1/2}e^{j2\pi fr}dH_{ij}(f) \tag{7-1-39}$$

其中 $$dH_{ij}(f)=E[dZ_i(f)dZ_j^{*}(f)]$$

称矩阵$H(f)\triangleq[H_{ij}(f)]$为积分谱矩阵，其对角线元素$H_{ii}(f)$为$\{x_{i,t}\}$的积分频谱，而$H_{ij}(f)$为$x_{i,t}$和$x_{j,t}$间的积分互频谱，如果每个$H_{ij}(f)$均可微，即存在$dH_{ij}(f)=G_{ij}(f)df$(所有$i,j$)，则式（7-1-39）可写成

$$R_{ij}(r)=\int_{-1/2}^{1/2}e^{j2\pi fr}G_{ij}(f)df \text{（所有}i,j\text{）} \tag{7-1-40}$$

其反演为

$$G_{ij}(f)=\sum_{r=-\infty}^{\infty}R_{ij}(r)e^{-j2\pi fr} \tag{7-1-41}$$

当$i=j$时$G_{ii}(f)$为$x_{i,t}$的（自）谱密度函数，而$i\neq j$时$G_{ij}(f)$为$x_{i,t}$和$x_{j,t}$的互谱密度函数。频谱矩阵为

$$G(f)=[G_{ij}(f)] \qquad i,j=1,\cdots,p \tag{7-1-42}$$

任何一对过程（$x_{i,t}$和$x_{j,t}$）的协谱、转象谱、互幅谱、相谱和相干谱等均可相应地加以定义。

式（7-1-40）和（7-1-41）的富氏变换关系可以写成更紧凑的形式如下

$$R(r)=\int_{-1/2}^{1/2} e^{i2\pi f r}G(f)df \tag{7-1-43}$$

$$G(f)=\sum_{r=-\infty}^{\infty} R(r)e^{-i2\pi f r} \tag{7-1-44}$$

此外，如果将多变量过程用向量$X_t=[x_{1,t},x_{2,t},\cdots,x_{p,t}]^T$表示，则可写出

$$R(r)=E[X_{t+r}X_t^*] \tag{7-1-45}$$

$$R^*(r)=R(-r) \tag{7-1-46}$$

（这里*号表示共轭转置），而过程的谱展式（7-1-38）写作

$$X_t=\int_{-1/2}^{1/2} e^{i2\pi f t}dZ(f) \tag{7-1-47}$$

其中　$dZ(f)=[dZ_1(f),dZ_2(f),\cdots,dZ_p(f)]^T$为列向量，谱矩阵$G(f)$可写成

$$G(f)df=E[dZ(f)dZ^*(f)] \tag{7-1-48}$$

可以看出

$$G^*(f)=G(f) \tag{7-1-49}$$

或

$$G_{ij}(f)=G_{ji}^*(f) \qquad 所有i,j \tag{7-1-50}$$

所以$G(f)$是埃尔米特（*Hermitian*）矩阵。还可证明它是正半定的，令$\lambda=[\lambda_1,\lambda_2,\cdots\lambda_p]$为任意（复）常数行向量，考虑单变量过程

$$y_t=\lambda X_t$$

其协方差函数为

$$R_y(r)=E[y_{t+\gamma}y_t^*]=\lambda E[X_{t+\gamma}X_t^*]\lambda^*=\lambda R(r)\lambda^*$$

则y_t的谱密度函数为

$$G_y(f)=\sum_{\gamma=-\infty}^{\infty}R_y(r)e^{-j2\pi f}=\lambda G(f)\lambda^*$$

由于对任意的$\lambda_1,\lambda_2,\cdots,\lambda_p$有$G_y(f)\geqslant0$（所有$f$），故$G(f)$为正半定矩阵。

7.2 具有线性关系的多变量过程谱分析

1. 严格线性关系 在2.4.4节中曾讨论过单变量线性模型"输入""输出"的谱之间的关系。今设y_t表"输入"，x_t表"输出"，它们满足严格线性关系（即不含其他噪声项），且不加物理可实现条件，即

$$x_t=\sum_{u=-\infty}^{\infty}g_u y_{t-u} \tag{7-2-1}$$

x_t和y_t的谱密度有如下关系（参见式（2-4-46））

$$G_x(f)=|\Phi(f)|^2G_y(f) \tag{7-2-2}$$

其中

$$\Phi(f)=\sum_{u=-\infty}^{\infty}g_u e^{-j2\pi fu} \tag{7-2-3}$$

是模型的传递函数，$\{g_u\}$为模型的格林函数。x_t具有有限方差的条件是

$$\int_{-1/2}^{1/2} |\Phi(f)|^2 G_y(f)df < \infty \tag{7-2-4}$$

和式（7-2-1）相似，对于q个输入过程$y_{1,t}, y_{2,t}, \cdots, y_{q,t}$和$p$个输出过程$x_{1,t}$, $x_{2,t}$, $\cdots$, $x_{p,t}$的多变量情况如图 7-4 所示。不同输入和输出之间有相互联系，且每个输出可以看成是所有输入过程线性函数之和。

对第i个输出来说

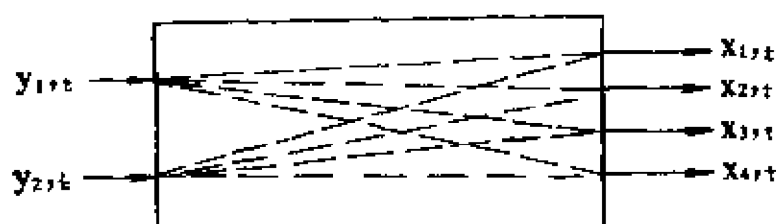


图7-4　多输入多输出系统

$$x_{i,t} = \sum_{u=-\infty}^{\infty} g_{i1}(u)y_{1,t-u} + \sum_{u=-\infty}^{\infty} g_{i2}(u)y_{2,t-u} + \cdots + \sum_{u=-\infty}^{\infty} g_{iq}(u)y_{q,t-u} \quad i=1,2,\cdots,p \tag{7-2-5}$$

或用矩阵表示

$$X_t = \sum_{u=-\infty}^{\infty} g(u)Y_{t-u} \tag{7-2-6}$$

其中　$X_t = [x_{1,t}\ x_{2,t} \cdots x_{p,t}]^T$

$Y_t = [y_{1,t}\ y_{2,t} \cdots y_{q,t}]^T$

$g(u) = [g_{ij}(u)]$　$u = 0, \pm 1, \pm 2, \cdots$为格林函数矩阵。

利用谱展式

$$x_{i,t}=\int_{-1/2}^{1/2} e^{i2\pi ft}dZ_i^{(x)}(f) \qquad i=1,2,\cdots,p$$

(7-2-7)

$$y_{j,t}=\int_{-1/2}^{1/2} e^{i2\pi ft}dZ_j^{(y)}(f) \qquad j=1,2,\cdots,q$$

(7-2-8)

可将式（7-2-5）右边第j项写成$\int_{-1/2}^{1/2} e^{i2\pi ft}\Phi_{ij}(f)dZ_j^{(y)}(f)$，其中

$$\Phi_{ij}(f)=\sum_{u=-\infty}^{\infty} g_{ij}(u)e^{-i2\pi f} \tag{7-2-9}$$

表示第j个输入和第i个输出之间的传递函数。由式（7-2-5）可以导出对每一个f有

$$dZ_i^{(x)}(f)=\Phi_{i1}(f)dZ_1^{(y)}(f)+\cdots+\Phi_{iq}(f)dZ_q^{(y)}(f)$$
$$i=1,\cdots,p \tag{7-2-10}$$

这个关系很重要，在时间域中（式（7-2-5））t时刻的第i个输出包含着所有输入的过去、现在和将来值的加权线性和，而在频率域中（式（7-2-10））关系比较简单，它实际上是一个典型的多元线性回归模型，而且如同单输入单输出情况一样，在频率f处输出的谱特性只取决于在同一频率f处的输入的谱特性。把式(7-2-10）用矩阵表示为

$$dZ^{(x)}(f)=\Phi(f)dZ^{(y)}(f) \tag{7-2-11}$$

其中$\Phi(f)=[\Phi_{ij}(f)]$称传递函数阵

$$\Phi(f)=\begin{bmatrix}\Phi_{11}(f)\Phi_{12}(f)\cdots\Phi_{1q}(f)\\ \Phi_{21}(f)\Phi_{22}(f)\cdots\Phi_{2q}(f)\\ \cdots\cdots\\ \Phi_{p1}(f)\Phi_{p2}(f)\cdots\Phi_{iq}(f)\end{bmatrix} \tag{7-2-12}$$

矩阵中第i行j列元素是第j输入对第i输出的传递函数。传递函数

阵完全地描述了多变量过程。

由式（7-2-11）还可求出输入和输出频谱矩阵间的关系。因

$$E[dZ^{(x)}(f)dZ^{(x)*}(f)]=\Phi(f)E[dZ^{(y)}(f)dZ^{(y)*}(f)]\Phi^*(f)$$

由式（7-1-47）得

$$G_x(f)=\Phi(f)G_y(f)\Phi^*(f) \tag{7-2-13}$$

它可看成是式（7-2-2）在多变量情况下的推广。由于$x_{i,t}$的方差是$G_x(f)$中第i个对角元素的积分，因此每个输出具有有限方差的条件（类似于式（7-2-4）为

$$tr\left[\int_{-1/2}^{1/2}\Phi(f)G_y(f)\Phi^*(f)df\right]<\infty \tag{7-2-14}$$

2. 附加噪声时的线性关系

（1）单输入单输出的系统。先讨论输出受噪声N_t扰动时的单输入单输出系统（图7-5），这时的关系为

$$x_t=\sum_{u=-\infty}^{\infty}g_u y_{t-u}+N_t \tag{7-2-15}$$

由于输出x_t的谱和y_t，N_t两者均有关，故式（7-2-2）不适用，因而不能只根据$G_x(f)$和$G_y(f)$来确定传递函数的模$|\Phi(f)|$，但若N_t和y_t无关，则可通过y_t和x_t的互谱$G_{xy}(f)$来确定出$\Phi(f)$。下面设N_t和y_t为无关的零均值实过程，这时

$$R_{xy}(r)=E[y_t x_{t+r}]=\sum_{u=-\infty}^{\infty}g_u E[y_t y_{t+r-u}]$$

$$=\sum_{u=-\infty}^{\infty}g_u R_{yy}(r-u) \tag{7-2-16}$$

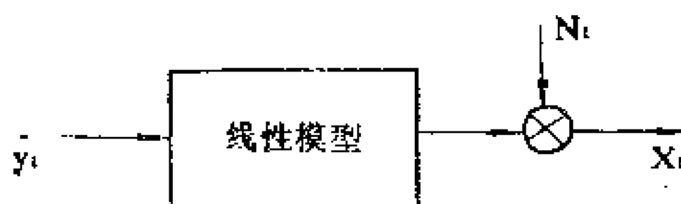


图7-5 带有输出噪声的单输入单输出系统

两边取富氏变换得

$$G_{xy}(f)=\sum_{\gamma=-\infty}^{\infty} e^{-j2\pi f\gamma}\left[\sum_{u=-\infty}^{\infty} g_u R_{yy}(r-u)\right]$$

$$=\sum_{u=-\infty}^{\infty} g_u e^{-j2\pi f u}\left[\sum_{\gamma=-\infty}^{\infty} e^{-j2\pi f(\gamma-u)} R_{yy}(r-u)\right]$$

$$=\Phi(f)G_{yy}(f) \qquad (7\text{-}2\text{-}17)$$

这里用$G_{yy}(f)$代替$G_y(f)$表示y_t的谱密度。于是得到线性模型的一个基本关系：

$$\Phi(f)=G_{xy}(f)/G_{yy}(f) \qquad (7\text{-}2\text{-}18)$$

即传递函数$\Phi(f)$等于输入输出互谱密度和输入谱密度之比。

式（7-2-18）也可利用谱展式直接导出，即先由式(7-2-15)写出

$$dZ_x(f)=\Phi(f)dZ_y(f)+dZ_N(f) \qquad (7\text{-}2\text{-}19)$$

两边乘以$dZ_y^*(f)$再取期望，考虑到$dZ_y(f)$ 和 $dZ_N(f)$ 无关，得

$$E[dZ_x(f)dZ_y^*(f)]=\Phi(f)E[|dZ_y(f)|^2] \qquad (7\text{-}2\text{-}20)$$

根据式（7-1-23）（7-1-21）便可导出式（7-2-18）。

从式（7-2-20）还可看到

$$\Phi(f)=\mathrm{Cov}[dZ_x(f),dZ_y(f)]/\mathrm{Var}[dZ_y(f)] \qquad (7\text{-}2\text{-}21)$$

这实际上是把式（7-2-19）看作是$dZ_x(f)$关于$dZ_y(f)$的线性回归时对$\Phi(f)$所作的最小二乘估计。换言之，如果给定两个过程x_t和y_t，则按式（7-2-18）求得的传递函数是对两者关系的最佳线性近似（在最小二乘意义下）。

y_t和x_t之间线性相依的程度可以用N_t的方差大小来度量，N_t的谱可由

$$G_{xx}(f)=|\Phi(f)|^2G_{yy}(f)+G_{NN}(f) \tag{7-2-22}$$

求出为

$$\begin{aligned}G_{NN}(f)&=G_{xx}(f)-|\Phi(f)|^2G_{yy}(f)\\&=\left[1-\frac{|G_{xy}(f)|^2}{G_{xx}(f)G_{yy}(f)}\right]G_{xx}(f)\\&=[1-|W_{xy}(f)|^2]G_{xx}(f)\end{aligned} \tag{7-2-23}$$

其中$|W_{xy}(f)|$是x_t和y_t在频率f处的相干。于是x_t机N_t的方差可表为

$$\sigma_x^2=\int_{-1/2}^{1/2}G_{xx}(f)\,|W_{xy}(f)|^2df+\sigma_N^2 \tag{7-2-24}$$

$$\sigma_N^2=\int_{-1/2}^{1/2}G_{NN}(f)df=\int_{-1/2}^{1/2}[1-|W_{xy}(f)|^2]G_{xx}(f)df \tag{7-2-25}$$

可见y_t和x_t间线性依赖程度（或σ_N^2的大小)取决于$|W_{xy}(f)|^2$在所有频率处对于1的接近程度，如果$|W_{xy}(f)|^2$在所有频率处均接近1，则两个序列可以得到很好的线性拟合，这也说明了$W_{xy}(f)$能够起频域相关系数的作用，由于它可通过x_t和y_t的谱特性求得，在N_t不可测的情况下用相干来说明线性关系是经常用到的。式（7-2-25）右边积分可以看成是最佳线性拟合情况下的残差方差值，因此它是拟合残差的下限。

从（7-2-23）～(7-2-25）诸式的推导中还可看出

$$\gamma^2 = |W_{xy}(f)|^2 = \frac{|G_{xy}|^2}{G_{xx}G_{yy}} = \frac{|\Phi|^2 G_{yy}}{G_{xx}}$$

$$= \frac{|\Phi|^2 G_{yy}}{|\Phi|^2 G_{yy} + G_{NN}}$$

$$= \frac{\text{归因于输入的输出功率}}{\text{总的输出功率}}$$

即γ^2是输出功率中归因于输入信号的比例数（分数），$\gamma^2 G_{xx}$是和输入有关的输出功率，而$(1-\gamma^2)G_{xx}$（即式(7-2-23)）是输出功率中的噪声成分（系统中可能的非线性会把部分输入信号转换成其他频率的能量，相干函数把这些能量当作噪声处理）。因此系统的信噪比可表为$\gamma^2/(1-\gamma^2)$，当然这不是整个信噪比，而是在某个频率处的信噪比，这个值对于分析结果是很有用的。γ^2可称为平方相干。如果在某频率处Φ或G_{yy}为零，则该频率处相干为零，全部输出功率都是噪声引起的。故相干小于1时至少有下列情况之一存在：1.量测受噪声污染，2.系统非线性，3.系统中有其他输入作用。

作为图7-5系统的图例，一个零均值单位方差的白噪声y_t通过AR(1)系统，而输出端附加了均值为零，方差0.1的白噪声(N_t)后构成输出x_t、图7-6a,b,c)分别表示y_t和x_t之间的互谱、相谱和平方相干，y_t和x_t的互相关是前面给出的图7-1曲线。

(2) 多输入多输出的系统。图7-7为有噪声时的多输入多输出系统，这时

$$X_t = \sum_{u=-\infty}^{\infty} g(u) y_{t-u} + N_t \qquad (7\text{-}2\text{-}26)$$

其中$N_t = [N_{1,t}, N_{2,t}, \cdots, N_{p,t}]^T$每一分量$N_{i,t}$和每个$y_{j,t}$均

$|\hat{G}_{xy}(f)|$

(a)

$\angle \hat{G}_{xy}(f)$

(b)

γ^2

(c)

图7-6 互谱、相谱和平方相干示例

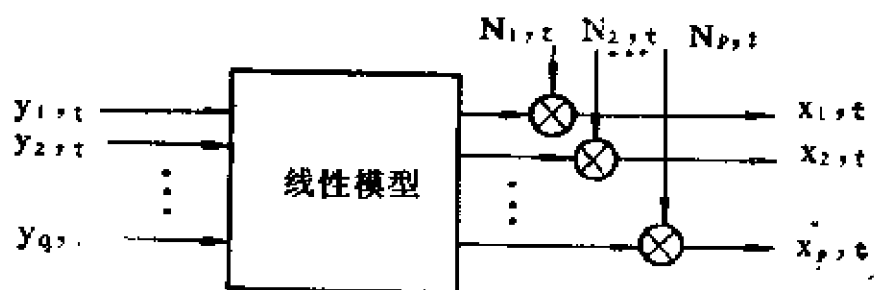


图7-7 带有输出噪声的多输入多输出系统

无关。此时式（7-2-11）变成

$$dZ^{(x)}(f)=\Phi(f)dZ^{(Y)}(f)+dZ^{(N)}(f) \qquad (7\text{-}2\text{-}27)$$

两边乘以$dZ^{(Y)*}(f)$并取期望，考虑到$dZ^{(Y)}(f)$和$dZ^{(N)}(f)$无关，可得

$$E[dZ^{(x)}(f)dZ^{(Y)*}(f)]=\Phi(f)E[dZ^{(Y)}(f)dZ^{(Y)*}(f)] \qquad (7\text{-}2\text{-}28)$$

或

$$G_{xy}(f)=\Phi(f)G_{yy}(f) \qquad (7\text{-}2\text{-}29)$$

其中$G_{yy}(f)$是Y_t的频谱矩阵（$q\times q$），$G_{xy}(f)$称为Y_t和X_t间的互谱矩阵（$p\times q$），它的第i,j元素是$x_{i,t}$和$y_{j,t}$间的互谱密度函数。若$G_{yy}(f)$非奇异，则有

$$\Phi(f)=G_{xy}(f)G_{yy}^{-1}(f) \qquad (7\text{-}2\text{-}30)$$

由式（7-2-27）有

$$G_{xx}(f)=\Phi(f)G_{yy}(f)\Phi^*(f)+G_{NN}(f)$$

$$=G_{xy}(f)G_{yy}^{-1}(f)G_{xy}^*(f)+G_{NN}(f) \qquad (7\text{-}2\text{-}31))$$

而残差过程N_t的谱密度为

$$G_{NN}(f)=G_{xx}(f)-G_{xy}(f)\ G_{yy}^{-1}(f)G_{xy}^*(f) \qquad (7\text{-}2\text{-}32)$$

和单输入单输出的情况类似，要估计在多大程度上能够用$y_{1,t},\cdots\cdots,y_{q,t}$的线性函数拟合某一个输出可用相干函数的概念，图7-8画出某一输出过程x_t（即$p=1$）和q个输入过程的关系：

$$x_t=\sum_{u=-\infty}^{\infty}g_1(u)y_{1,t-u}+\cdots\cdots+\sum_{u=-\infty}^{\infty}g_q(u)y_{q,t-u}+N_t \qquad (7\text{-}2\text{-}33)$$

我们定义在某频率f处的“多重平方相干”为归因于$y_{1,t},\cdots\cdots,$

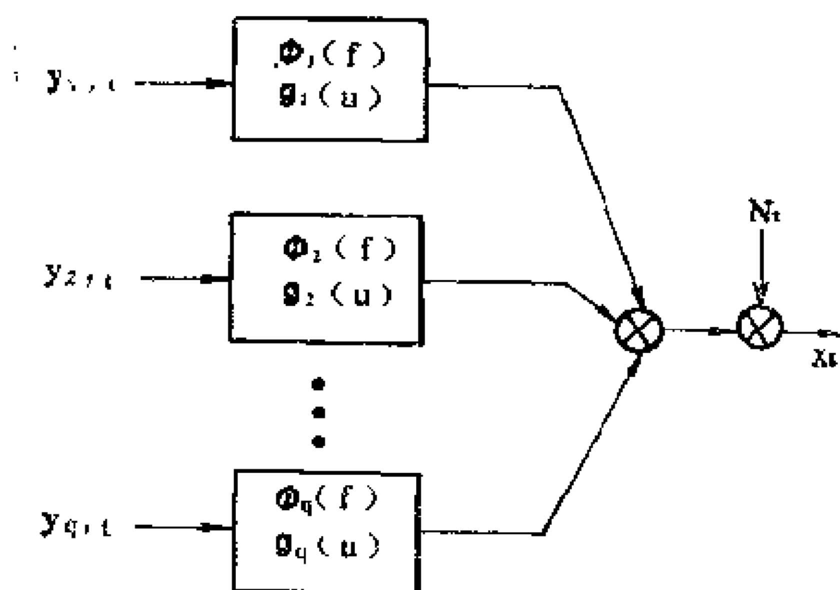


图7-8 带有噪声的多输入单输出系统

$y_{q,t}$的那一部分输出功率，由式（7-2-32）写出

$$G_{NN}(f)=G_{xx}(f)-G_{xy}(f)G_{yy}^{-1}(f)G_{xy}^{*}(f)$$

$$=G_{xx}(f)\left[1-W_{xy_1\cdots y_q}^{2}(f)\right] \qquad (7\text{-}2\text{-}34)$$

其中

$$W_{xy_1\cdots y_q}^{2}(f)\equiv\frac{G_{xy}(f)G_{yy}^{-1}(f)G_{xy}^{*}(f)}{G_{xx}(f)} \qquad (7\text{-}2\text{-}35)$$

就称为频率f处x_t和$y_{1,t}$，……，$y_{q,t}$之间的多重平方相干。由于这里x_t是标量，故$G_{xy}(f)$为$1\times q$阵，$G_{yy}(f)$为$q\times q$阵，$G_{xy}^{*}(f)$为$q\times 1$阵。式（7-2-34）可看作是单输入时式（7-2-23）的变形，如果在所有f处$W_{xy_1\cdots y_q}^{2}(f)=1$，则$x_t$和$y_{1,t}$，…，$y_{q,t}$之间是精确的线性关系。

然而在多输入的情况下还有一个新问题，譬如人们可能要知道两个变量（如x_t和$y_{1,t}$）所具有的相干性是由于它们本身固有的还是由于受别的变量（如$y_{2,t}$）对x_t和$y_{1,t}$的共同影响而产生

的。为此，要考察除去$y_{2,t}$对x_t和$y_{1,t}$的影响之后新过程的相干性一即排除$y_{2,t}$时x_t和$y_{1,t}$间的“偏相干”。将这种新的过程记作

$$\eta_{1,t}=x_t-\sum_{u=-\infty}^{\infty}b_1(u)y_{2,t-u}$$

$$\eta_{2,t}=y_{1,t}-\sum_{u=-\infty}^{\infty}b_2(u)y_{2,t-u}$$

其中 $\{b_1(u)\}$，$\{b_2(u)\}$ 分别根据$E[\eta_{1,t}^2]$ 和 $E[\eta_{2,t}^2]$ 为最小来确定，由式（7-2-18）（7-2-21）知，它们应当有

$$B_1(f)=\sum_{u=-\infty}^{\infty}b_1(u)e^{-j2\pi fu}=G_{xy_2}(f)/G_{y_2y_2}(f)$$

$$B_2(f)=\sum_{u=-\infty}^{\infty}b_2(u)e^{-j2\pi fu}=G_{y_1y_1}(f)/G_{y_2y_2}(f)$$

由$\eta_{1,t}$和$\eta_{2,t}$的谱展式写出

$$\begin{aligned}dZ_{\eta_1}(f)&=dZ_x(f)-B_1(f)dZ_{y_2}(f)\\ dZ_{\eta_2}(f)&=dZ_{y_1}(f)-B_2(f)dZ_{y_2}(f)\end{aligned}\qquad(7\text{-}2\text{-}36)$$

由$E[dZ_{\eta_1}(f)dZ^*_{\eta_2}(f)]$得$\eta_{1,t}$和$\eta_{2,t}$的互谱密度为

$$\begin{aligned}G_{\eta_1\eta_2}(f)&=G_{xy_1}(f)-B_1(f)G_{y_2y_1}(f)-B_2^*(f)G_{xy_2}(f)\\ &\quad+B_1(f)B_2^*(f)G_{y_2y_2}(f)\\ &=G_{xy_1}(f)-\frac{G_{xy_2}(f)G_{y_2y_1}(f)}{G_{y_2y_2}(f)}\end{aligned}\qquad(7\text{-}2\text{-}37)$$

称其为排除$y_{2,t}$时x_t和$y_{1,t}$的偏互谱密度函数，并记作$G_{xy_1\cdot y_2}(f)$。$\eta_{1,t}$和$\eta_{2,t}$的（复）相干则定义为偏（复）相干$W_{xy_1\cdot y_2}(f)$，即

$$W_{xy_1\cdot y_2}(f)=G_{\eta_1\eta_2}(f)/[G_{\eta_1\eta_1}(f)G_{\eta_2\eta_2}(f)]^{1/2} \tag{7-2-38}$$

而利用式（7-2-36）不难得出

$$G_{\eta_1\eta_1}(f)=G_{xx}(f)[1-|W_{xy_2}(f)|^2]$$
$$G_{\eta_2\eta_2}(f)=G_{y_1y_1}(f)[1-|W_{y_2y_1}(f)|^2] \tag{7-2-39}$$

将式（7-2-39）（7-2-37）代入（7-2-38）最后得

$$W_{xy_1\cdot y_2}(f)=\frac{W_{xy_1}(f)-W_{xy_2}(f)W_{y_2y_1}(f)}{[(1-|W_{xy_2}(f)|^2)(1-|W_{y_2y_1}(f)|^2)]^{1/2}} \tag{7-2-40}$$

在更一般的情况下偏相干的表达式是考虑排除几个过程的影响时两个过程的相干性。如果给定任意多变量平稳过程$x_t=[x_{1,t},x_{2,t},\cdots,x_{p,t}]^T$，取任一$x_{i,t}$作为“输出”，其余的$[x_{j,t},x_{k,t},x_{l,t},x_{m,t},\cdots]$作为“输入”，可定义并求出当排除任意子集$[x_{k,t},x_{l,t},x_{m,t},\cdots]$时任何一对$x_{i,t}$和$x_{j,t}$之间的偏相干$W_{ij\cdot klm}(f)$。

7.3 互谱特性和相干函数的估计和应用

首先来讨论根据有限的实际记录数据估计自谱和互谱特性，这时已知的是$t=1,2,\cdots,N$时实值多变量平稳过程$X_t=[x_{1,t},x_{2,t},\cdots,x_{p,t}]^T$的N个观察值。对于$x_{i,t}$的（自）谱密度$G_{ii}(f)$的估计显然可以用第四章所介绍的方法，而对于$x_{i,t}$和$x_{j,t}$的互谱特性$G_{ij}(f)$的估计方法也和以前大体类似。

如果从相关函数出发，则首先是建立互协方差函数$R_{ij}(\tau)$的估计。

$$\hat{R}_{ij}(\tau)=\frac{1}{N}\sum_t(x_{i,t}-\bar{x}_j)(x_{j,t+\tau}-\bar{x}_i)$$

$$r=0,\pm1,\cdots,\pm(N-1) \qquad (7\text{-}3\text{-}1)$$

这里当$r\geqslant0$时求和是从$t=1$到$N-r$，当$r<0$时是从$t=1-r$到N。$\bar{x}_j$和$\bar{x}_i$分别是$x_{j,t}$和$x_{i,t}$的样本均值，即$\bar{x}_i=\frac{1}{N}\sum_{t=1}^{N}x_{i,t}$。不难看出$\hat{R}_{ij}(r)$是$R_{ij}(r)$的渐近无偏估计

$$E[\hat{R}_{ij}(r)]=R_{ij}(r)+O(1/N)$$

而且和3.1.2节讨论类似，可以证明当X_t为正态过程且N很大时有

$$\text{Cov}[\hat{R}_{ij}(r),\hat{R}_{ij}(r+v)]\approx\frac{1}{N}\sum_{m=-\infty}^{\infty}[R_{ii}(m)R_{jj}(m+v)$$

$$+R_{ij}(m+r+v)R_{ji}(m-r)] \qquad (7\text{-}3\text{-}2)$$

互谱密度函数的估计式为

$$\hat{G}_{ij}(f)=\sum_{r=-(N-1)}^{N-1}q_N(r)\hat{R}_{ij}(r)e^{-j2\pi fr} \qquad (7\text{-}3\text{-}3)$$

式中$q_N(r)$为协方差迟后窗。显然，当$i=j$时式（7-3-3）就是自谱的估计式。但应指出，对于各个不同过程的自谱$G_{ii}(f)$或互谱$G_{ij}(f)$的估计没有理由都要采用同样的迟后窗，因为$q_N(r)$的选择要和$R_{ii}(r)$的衰减快慢相“匹配”，或等效地说，谱窗$Q_N(f)$也要和$G_{ii}(f)$的带宽相“匹配”。在一般多变量过程中，不同分量自谱的带宽不见得相等或接近，互谱的带宽也不一定和自谱的带宽相当，因此即使我们在所有的谱估计式中都取相同数学表达式的$q_N(r)$，在计算各个自谱和互谱的估计时截取宽度可以取不同的值。换句话说，式（7-3-3）中的$q_N(r)$还和i,j有关。故准确地应当写成$q_N^{(i,j)}(r)$，不过为了简便我们仍保持原来的符号，并假设采用相同的迟后窗。这样当用谱矩阵形式表示时，式(7-3-3)可以写成更紧凑的形式

$$\hat{G}(f)=\sum_{\tau=-(N-1)}^{N-1} q_N(\tau)\hat{R}(\tau)e^{-j2\pi f\tau} \tag{7-3-4}$$

其中$\hat{R}(\tau)=[\hat{R}_{ij}(\tau)]$是迟后为$\tau$的样本协方差阵。有关多变量谱估计的分析同单变量的情况很类似，许多多变量的公式都相当于单变量情况的变形，推导也基本相同，这里不再细述。

式（7-3-3）在实际计算中往往是利用式（7-1-28）和（7-1-29）先计算$\hat{R}_{ij}(\tau)+\hat{R}_{ij}(-\tau)$和$\hat{R}_{ij}(\tau)-\hat{R}_{ij}(-\tau)$，作协谱和重谱的估计

$$\hat{P}_{ij}(f)=Re[\hat{G}_{ij}(f)]=\frac{1}{2}\sum_{\tau=-(N-1)}^{N-1} q_N(\tau)[\hat{R}_{ij}(\tau)+\hat{R}_{ij}(-\tau)]\cos 2\pi f\tau$$

$$\hat{Q}_{ij}(f)=-Im[\hat{G}_{ij}(f)]=\frac{1}{2}\sum_{\tau=-(N-1)}^{N-1} q_N(\tau)[\hat{R}_{ij}(\tau)-\hat{R}_{ij}(-\tau)]\sin 2\pi f\tau$$

于是互幅谱$M_{ij}(f)$和相谱$\psi_{ij}(f)$的估计为

$$\hat{M}_{ij}(f)=|\hat{G}_{ij}(f)|=[\hat{P}_{ij}^2(f)+\hat{Q}_{ij}^2(f)]^{1/2}$$

$$\hat{\psi}_{ij}(f)=\arg\hat{G}_{ij}(f)=\tan^{-1}[-\hat{Q}_{ij}(f)/\hat{P}_{ij}(f)]$$

$x_{i,t}$和$y_{j,t}$之间的传递函数估计为

$$\hat{\Phi}_{ij}(f)=[\hat{P}_{ij}(f)-j\hat{Q}_{ij}(f)]/\hat{G}_{ii}(f) \tag{7-3-5}$$

$\hat{\Phi}_{ij}(f)$的幅角和$\hat{\psi}_{ij}(f)$相同，而模为

$$|\hat{\Phi}_{ij}(f)|=[\hat{P}_{ij}^2(f)+\hat{Q}_{ij}^2(f)]^{1/2}/\hat{G}_{ii}(f)=\hat{M}_{ij}(f)/\hat{G}_{ii}(f) \tag{7-3-6}$$

将互谱估计代入式（7-1-33）可得相干的估计

$$|\hat{W}_{ij}(f)|=\frac{|\hat{G}_{ij}(f)|}{[\hat{G}_{ii}(f)\hat{G}_{jj}(f)]^{1/2}}=\left[\frac{\hat{P}_{ij}^2(f)+\hat{Q}_{ij}^2(f)}{\hat{G}_{ii}(f)\hat{G}_{jj}(f)}\right]^{1/2} \tag{7-3-7}$$

同单变量谱估计的直接法类似，在多变量谱估计中有（修正的）周期图矩阵$\mathbf{I}_N(f)$，它定义为

$$\mathbf{I}_N(f)=[I_{N,ij}(f)] \qquad i,j=1,2,\cdots,p \tag{7-3-8}$$

其中

$$I_{N,ij}(f)=\frac{1}{N}\left(\sum_{t=1}^{N}x_{i,t}e^{-j2\pi ft}\right)\left(\sum_{t=1}^{N}x_{j,t}e^{-j2\pi ft}\right)^{*} \tag{7-3-9}$$

当$i=j$时$I_{N,ii}$就是$x_{i,t}$的周期图，$i\neq j$时$I_{N,ij}$是$x_{i,t}$和$x_{j,t}$间的互周期图。为便于计算，互周期图写成

$$I_{N,ij}(f)=[A_{x_i}(f)A_{x_j}(f)+B_{x_i}(f)B_{x_j}(f)]-j[B_{x_i}(f)A_{x_j}(f)-A_{x_i}(f)B_{x_j}(f)] \tag{7-3-10}$$

其中 $$A_{x_i}(f)=\frac{1}{\sqrt{N}}\sum_{t=1}^{N}x_{i,t}\cos2\pi ft$$

$$B_{x_i}(f)=\frac{1}{\sqrt{N}}\sum_{t=1}^{N}x_{i,t}\sin2\pi ft$$

如同$I_{N,ii}(f)$代表$x_{i,t}$自谱的原始样本（未经平滑），$I_{N,ij}(f)$代表$x_{i,t}$和$x_{j,t}$间原始的样本互谱函数。

为了减小周期图谱估计的方差，可以采用和单变量中类似的分段周期图的总体平均方法，不过这里除了对模以外还要对相位进行总体平均，这些量利用各段周期图的实部和虚部都不难计算。

关于互谱的应用，前面曾经提到，由于实际情况下系统内部或外部的噪声及非线性等因素都会影响“输入”和“输出”间的线性关系，这时如果用$\hat{G}_x/\hat{G}_y$来估计系统的传递函数将是有偏的，因为$\hat{G}_x=|\Phi|^2\hat{G}_y+\hat{G}_N$，它和$\hat{G}_y$的比值不是$|\Phi|^2$，而且无论对样本估计作任何平滑都不会使比值收敛到$|\Phi|^2$的真值，只有利用互谱即$\hat{G}_{xy}/\hat{G}_{yy}$才能得到$\Phi$的无偏估计，因为$\hat{G}_{xy}=\Phi\hat{G}_{yy}+\hat{G}_{Nx}$，当$N_t$和$x_t$无关的情况下其互谱估计通过平均的方

法可以趋于零从而得出真正的传递函数。当然，问题是需要多少个样本周期图的平均才能得到希望的准确结果呢？这个平均数目的确定将要用到相干函数。

这里提到的准确性指的是统计上的精度，可以用例如“90%的置信限”来给这个精度定量，贝达特(*Bendat*)，皮尔索(*Piersol*)和古德曼（*Goodman*）都曾讨论并证明了单输入和多输入情况下某个传递函数Φ_i在频率f处的$(1-\alpha)\times100\%$的置信区间为（参见〔21〕〔10〕）

$$|\hat{\Phi}_i(f)-\Phi_i(f)|^2\leqslant r^2(f) \tag{7-3-11}$$

其中$\hat{\Phi}_i$是传递函数的样本估计（由样本互谱估计得），$r(f)$和谱估计、相干估计、估计的自由度以及显著水平有关，在单输入情况下

$$r^2(f)=\frac{2}{\nu-2}F_{2,\nu-2;\alpha}(1-\hat{W}^2_{xy})\hat{G}_x(f)/\hat{G}_y(f) \tag{7-3-12}$$

其中ν为谱估计的自由度（当分线L段作平均时$\nu=2L$），

$F_{2,\nu-2;\alpha}$是自由度为$\nu_1=2,\nu_2=\nu-2$的F分布100α百分比点，

$\hat{W}_{xy}$是输入y_t和输出x_t间的相干估计，

$\hat{G}_x$和$\hat{G}_y$分别为x_t和y_t的功率谱估计。

例如某单输入系统在给定频率f_0处有$|\hat{\Phi}(f_0)|=1.25$，另根据谱密度的估计有$\hat{G}_x(f_0)=0.28$，$\hat{G}_y(f_0)=0.13$，$\hat{W}_{xy}(f_0)=0.7$，，估计是在50次平均的情况下取得的，这时真正$|\Phi(f_0)|$的90%置信区间可计算如下：

由于$L=50$故$\nu=100$，由F分布表查得$F_{2,98;0.1}=2.36$，按式（7-3-12）得

$$r^2(f_0)=\frac{2}{98}\times 2.36(1-0.7)\left(\frac{0.28}{0.13}\right)=0.031$$

$$r(f_0)=0.176$$

于是有$1.074\leqslant|\Phi(f_0)|\leqslant 1.426$，增加平均次数可进一步减小置信区间。

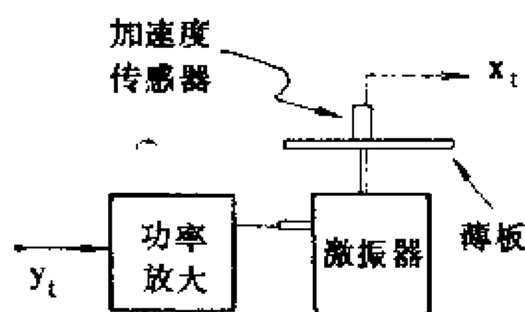
实用中经常在传递函数估计的同时也计算出相干并加以对照。图7-9a所安排的实验可以说明传递函数估计会受其他非外加输入的信号所"污染"，并且可以看出相干函数能够发现这种信的号存在。

这里要测试一块薄板的机械惯性（它等于加速度/力），用伪随机信号使激振器产生机械力（输入），加速度传感器测量加速度（输出），其传递函数就代表惯性。图7-9b为其曲线，而且还给出相干函数曲线。除$f=0$Hz处以外（这里加速度传感器输出为零）相干函数几乎处处为1，这说明量测条件良好（从信噪比看）。只是在100Hz处相干函数减到0.83，表明信噪比为0.83（1—0.83）或7 db，这是由于该频率处加速度传感器输出较低。

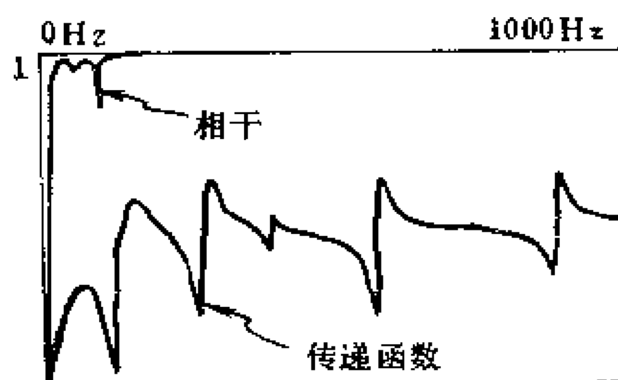
如果把条件稍加改变，将一个小螺丝松动地放在板上的孔中，螺丝的质量使谐振频率稍有下降，而由于它的松动对板产生撞击如同外加作用，这种做法实际上是在系统中构造了一种非线性元件，它把一部分输入能量转换到其他频率（由图7-9c可见其相应曲线），这种情况主要出现在输出响应较弱的地方，而相干函数大部分还是1。还可看出在260Hz处的影响很小，这表明螺丝振动所产生的能量在频率范围内并非均匀分布。

显然，在估计传递函数时辅以相干估计可以提示我们注意存在着的信噪比问题，并给出有关其频率和大小的信息。

另一种应用问题是不要求"输入"和"输出"间传递函数的精确形式，而是要分析因果关系，即确定各种因素对所观察的x。



（a）



（b）



（c）

图7-9 传递函数测试原理和相应曲线

的影响程度如何。例如要判断车间中哪一台机器的噪声对某一指定点的影响最大，哪一台机器的影响最小。传统的方法是根据假定的噪音源和输出信号x_t之间的互相关来确定。相干函数显然也能找到它们的因果关系，而且比起互相关方法还有另外的好处。因为我们知道，互相关是时间函数，其最大值所对应的时间是噪音源与观察点之间的传播时间滞后，而相干函数是频率函数，其最大值是出现在能量传输最大的频率处。根据这种频率分布情况可以设法削弱这种影响（如噪声、振动等），因此相干函数不仅揭示了因果关系的程度，而且对选择解决问题的办法有直接的指导意义。

作为因果关系的度量，希望得到统计精度较高的相干估计，为此在计算过程中要进行某种平均以减小方差。安诺克森（*Enochson*）和古德曼（*Goodman*）证明了相干的下述变换近似服从正态分布，该变换取为

$$\omega(f)=\frac{1}{2}\ln\left[\frac{1+\hat{W}_{xy}(f)}{1-\hat{W}_{xy}(f)}\right]=\tanh^{-1}\hat{W}_{xy}(f)$$

它的正态分布所具有的均值和方差分别为

$$\mu_\omega(f)=(\nu-2)^{-1}+\tanh^{-1}W_{xy}(f)$$

$$\sigma_\omega^2=(\nu-2)^{-1}$$

利用标准化正态分布的关系

$$P\left[Z_{1-\alpha/2}<\frac{\omega(f)-\mu_\omega(f)}{\sigma_\omega}\leqslant Z_{\alpha/2}\right]=1-\alpha$$

（Z_α为100α百分比点），经过整理可将该式表为

$$\tanh[\omega(f)-(\nu-2)^{-1}-\sigma_\omega Z_{\alpha/2}]<W_{xy}(f)\leqslant\tanh[\omega(f)-(\nu-2)^{-1}+\sigma_\omega Z_{\alpha/2}]$$

这就是$W_{xy}(f)$的$(1-\alpha)\times100\%$置信区间，它是ν、$\hat{W}_{xy}(f)$和α的函数，平方相干（γ^2）的置信限是$W_{xy}(f)$相应界限的平方。

表7-1列出γ^2的90%置信限（上限和下限的值），$\hat{W}_{xy}(f)$的计算通常是在计算机或专用的数字分析仪中进行的，开始时可取较少的平均数目，根据求得的$\hat{W}_{xy}$可查表得到置信区间，如果不够满意可增加平均数目重新计算。

表7-1　相干函数估计的90%置信限

（表中数值为γ^2上限和γ^2下限）

$\hat{\gamma}^2$ ＼ 平均数目 L	16	32	64	128	256
0.4	.15, .59	.23, .54	.28, .50	.32, .47	.34, .45
0.5	.25, .67	.33, .63	.39, .59	.42, .57	.45, .55
0.6	.36, .74	.45, .71	.50, .68	.53, .66	.55, .64
0.7	.50, .81	.57, .78	.61, .76	.64, .75	.66, .73
0.8	.65, .88	.70, .86	.74, .84	.76, .83	.77, .82
0.9	.81, .94	.85, .93	.87, .92	.88, .92	.88, .91

下面是一个实验例子，它有两个输入作用到系统后组合成为总的输出信号，通过本例说明如何利用相干函数区分出两个输入各自的效果。

有时用简单的谱分析便可以找出因果关系，如输入中的某周期分量会在具有因果关系的输出中出现，但是当谱是连续的情况下（如随机过程）就没有什么输入输出共同具有的某分量的谱线，图7-10a的情形就是如此，其中$y_{1,t}$和$y_{2,t}$是作为输入的两个相同的噪声源，这里“相同”是指它们的功率谱都是平坦的且高度一样，因而靠谱分析是无法区分的，虚线框中所示系统系两个简单的高通和低通滤波器，信号分别通过后线性叠加产生输出x_t，调节滤波器参数使x_t的谱也是平坦的。

如果只能取出$y_{1,t}$，$y_{2,t}$和x_t而且不能切断噪声源，如何才能确定每个输入对输出的因果关系？或者说，如何才能把上述情况（被合成的两个过程的频率分布是不同的）和频率分布相同的

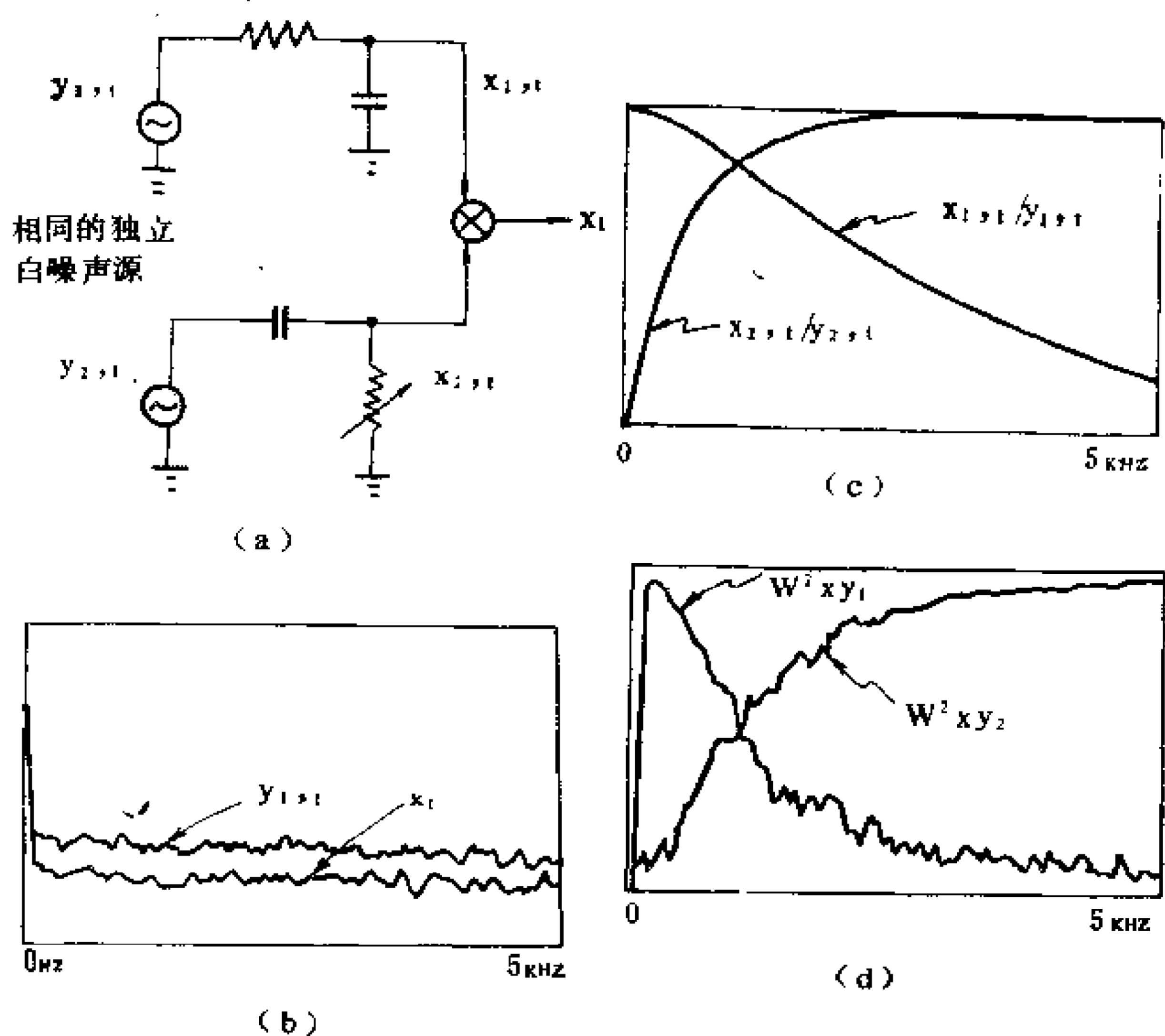


图7-10　因果关系测试示例

合成情况区分开来呢？图7-10*d*所示的相干估计曲线可以作出回答。因为$y_{1,t}$和x_t的相干函数在低频部分接近1，所以在低频段$y_{1,t}$在输出中占主要份量，至于高频段则以$y_{2,t}$为主。在任一频率处两个平方相干之和为1，这和它们的功率分数的意义是一致的。还可以指出：如果系统不使两个分量（$x_{1,t}$和$x_{2,t}$）的频率分布有差别（即均为平坦谱），则上述两个相干函数也是平坦的，其高度代表每个输入对输出功率的贡献。

7.4 多变量ARMA序列. 多变量AR模型的辨识

1. 多变量过程的矩阵差分方程表示 第二章中讨论过的三种单变量过程的模型，即AR、MA和$ARMA$都可以相应地推广到多变量过程，这主要就是把单变量模型中的标量参数用矩阵参数代替，因此它们是属于矩阵差分方程模型。

p变量的$AR(n)$模型可写成

$$X_t+A_1X_{t-1}+\cdots+A_nX_{t-n}=\varepsilon_t \tag{7-4-1}$$

其中$X_t=[x_{1t},\cdots,x_{pt}]^T$为$p$变量零均值平稳随机序列，$A_1,\cdots,A_n$都是$p\times p$阵。$\varepsilon_t=[\varepsilon_{1t},\cdots,\varepsilon_{pt}]^T$是$p$变量的白噪声序列，满足

$$E[\varepsilon_t]=0,\quad E[\varepsilon_t\varepsilon_s^T]=\begin{cases}Q & t=s\\ 0 & t\neq s\end{cases} \tag{7-4-2}$$

这里Q是p维正定阵，ε_t的各分量都是单变量白噪声过程，在不同时刻彼此无关，但在同一时刻，它们可能是互相相关的，因此Q在一般情况下是方差协方差阵——不一定是主对角阵。定义$A(B)$为下列矩阵多项式（B为后移算子）

$$A(B)=\sum_{i=0}^{n}A_iB^i\qquad(A_0=I) \tag{7-4-3}$$

则多变量过程的平稳条件为：$|A(B)|$的全部零点在单位园外。虽然多变量AR模型的矩阵形式和单变量形式极其相似，但多变量模型在结构上要复杂得多。应当注意，如果式（7-4-1）写成标量形式，那么每个对应于$x_{i,t}$的方程不仅包含$x_{i,t}$的迟后值，而且也包含所有其它变量$\{x_{j,t}\}$的迟后值。例如二变量$AR(2)$模型的标量形式为

$$x_{1t}+\theta_{11}x_{1,t-1}+\theta_{12}x_{2,t-1}+\varphi_{11}x_{1,t-2}+\varphi_{12}x_{2,t-2}=\varepsilon_{1t}$$

$$x_{2t}+\theta_{21}x_{1,t-1}+\theta_{22}x_{2,t-1}+\varphi_{21}x_{1,t-2}+\varphi_{22}x_{2,t-2}=\varepsilon_{2t}$$

这里

$$A_1=\begin{pmatrix}\theta_{11} & \theta_{12}\\ \theta_{21} & \theta_{22}\end{pmatrix},\qquad A_2=\begin{pmatrix}\varphi_{11} & \varphi_{12}\\ \varphi_{21} & \varphi_{22}\end{pmatrix}$$

p变量的MA(m)模型的矩阵形式为

$$X_t=\boldsymbol{\varepsilon}_t+B_1\boldsymbol{\varepsilon}_{t-1}+\cdots+B_m\boldsymbol{\varepsilon}_{t-m} \tag{7-4-4}$$

定义

$$B(B)=\sum_{i=0}^{m}B_iB^i\quad(B_0=I) \tag{7-4-5}$$

MA模型的可逆条件为 $|B(B)|$ 的全部零点在单位圆外。以二变量$MA(2)$模型为例，写出它的标量形式为

$$x_{1t}=\varepsilon_{1t}+\gamma_{11}\varepsilon_{1,t-1}+\gamma_{12}\varepsilon_{2,t-1}+\delta_{11}\varepsilon_{1,t-2}+\delta_{12}\varepsilon_{2,t-2}$$

$$x_{2t}=\varepsilon_{2t}+\gamma_{21}\varepsilon_{1,t-1}+\gamma_{22}\varepsilon_{2,t-1}+\delta_{21}\varepsilon_{1,t-2}+\delta_{22}\varepsilon_{2,t-2}$$

这里

$$B_1=\begin{pmatrix}\gamma_{11} & \gamma_{12}\\ \gamma_{21} & \gamma_{22}\end{pmatrix},\qquad B_2=\begin{pmatrix}\delta_{11} & \delta_{12}\\ \delta_{21} & \delta_{22}\end{pmatrix}$$

p变量的ARMA模型为

$$X_t+A_1X_{t-1}+\cdots+A_nX_{t-n}=\boldsymbol{\varepsilon}_t+B_1\boldsymbol{\varepsilon}_{t-1}+\cdots+B_m\boldsymbol{\varepsilon}_{t-m}$$

其算子形式为

$$A(B)X_t=B(B)\boldsymbol{\varepsilon}_t \tag{7-4-7}$$

这里$A(B)$和$B(B)$按式（7-4-3）和（7-4-5）定义。ARMA模型平稳和可逆的条件是 $|A(B)|$ 和 $|B(B)|$ 的零点都在单位圆外。一个二变量ARMA(n,m)模型的标量形式为

$$\begin{aligned}x_{1t}&+a_{11}x_{1,t-1}+a_{12}x_{2,t-1}+\cdots+a_{n1}x_{1,t-n}+a_{n2}x_{2,t-n}\\&=\varepsilon_{1t}+b_{11}\varepsilon_{1,t-1}+b_{12}\varepsilon_{2,t-1}+\cdots+\\&\quad b_{m1}\varepsilon_{1,t-m}+b_{m2}\varepsilon_{2,t-m}\end{aligned}$$

$$x_{2t}+a'_{11}x_{1,t-1}+a'_{12}x_{2,t-1}+\cdots+a'_{n1}x_{1,t-n}+a'_{n2}x_{2,t-n}$$
$$=\varepsilon_{2,t}+b'_{11}\varepsilon_{1,t-1}+b'_{12}\varepsilon_{2,t-1}+\cdots+$$
$$b'_{m1}\varepsilon_{1,t-m}+b'_{m2}\varepsilon_{2,t-m}$$

和单变量类似，平稳多变量模型可以写成传递形式

$$X_t=\sum_{i=0}^{\infty}g(i)\varepsilon_{t-i} \tag{7-4-8}$$

令

$$G(B)=\sum_{i=0}^{\infty}g(i)B^i \tag{7-4-9}$$

则式（7-4-8）亦可写成

$$X_t=G(B)\varepsilon_t$$

显然对于$AR(n)$模型

$$G(B)=[A(B)]^{-1}$$

对于$MA(m)$模型

$$G(B)=B(B)$$

而对于ARMA（n,m）模型

$$G(B)=[A(B)]^{-1}B(B)$$

对于这些模型来说，矩阵多项式均为有理形式。

2. 多变量自回归模型的参数估计 对p变量随机序列建立多变量自回归模型，就是由观察值$\{X_t\}$估计出多变量AR模型的阶数n，估计出系数矩阵$A_j(j=1,2,\cdots n)$及平稳白噪声的方差阵Q。关于如何定阶的问题将在后面讨论，现在假定阶数n为已知，求$A_j(j=1,2,\cdots n)$和Q的估计值。

(1) 系数矩阵的耶尔一瓦克尔估计。设$\{X_t\}$为多变量$AR(n)$序列，其模型为（后面以φ代A）

$$X_t-\varphi_1x_{t-1}-\cdots-\varphi_nX_{t-n}=\varepsilon_t \tag{7-4-10}$$

式（7-4-10）两边右乘以X^T_{t-r}并求数学期望，得

$$R(r)=\sum_{j=1}^{n}\varphi_j R(r-j) \qquad r=1,2,\cdots \tag{7-4-11}$$

$$R(0)=\sum_{i,\ j=1}^{n}\varphi_j R(i-j)\varphi_i^{\mathrm{T}}+Q$$

其中$R(r)$，$r=0,1,2,\cdots,n$都是$p\times p$矩阵。利用$R(-r)=R(r)^{\mathrm{T}}$，由（7-4-10）可得下面矩阵线性方程组（现考虑实过程）

$$\begin{pmatrix} R(1)^{\mathrm{T}} \\ R(2)^{\mathrm{T}} \\ \vdots \\ R(n)^{\mathrm{T}} \end{pmatrix}=\begin{pmatrix} R(0) & R(1)\cdots & R(n-1) \\ R(1)^{\mathrm{T}} & R(0)\cdots & R(n-2) \\ \cdots & & \\ R(n-1)^{\mathrm{T}} & R(n-2)^{\mathrm{T}}\cdots & R(0) \end{pmatrix}\begin{pmatrix} \varphi_1^{\mathrm{T}} \\ \varphi_2^{\mathrm{T}} \\ \vdots \\ \varphi_n^{\mathrm{T}} \end{pmatrix} \tag{7-4-12}$$

在线性方程组（7-4-12）中，当数据长度N充分大时可取

$$\hat{R}(r)=\frac{1}{N}\sum_{t=r+1}^{N}X_t X_{t-r}^{\mathrm{T}} \tag{7-4-13}$$

作为$R(r)$的估计值。由于$\hat{R}(r)$，$r=0,1,2,\cdots,n$是正定阵，故在式（7-4-12）中用$\hat{R}(r)$代替$R(r)$，它右边的系数矩阵为可逆，令

$$\Gamma_n=\begin{pmatrix} \hat{R}(0) & \hat{R}(1) & \cdots\hat{R}(n-1) \\ \hat{R}(1)^{\mathrm{T}} & \hat{R}(0) & \cdots\hat{R}(n-2) \\ \vdots & & \\ \hat{R}(n-1)^{\mathrm{T}} & \hat{R}(n-2)^{\mathrm{T}} & \cdots\hat{R}(0) \end{pmatrix},$$

$$\hat{\eta}_n=\begin{pmatrix} \hat{R}(1)^{\mathrm{T}} \\ \hat{R}(2)^{\mathrm{T}} \\ \vdots \\ \hat{R}(n)^{\mathrm{T}} \end{pmatrix}$$

$$\hat{\Phi}_n=\begin{pmatrix} \hat{\varphi}_1^{\mathrm{T}} \\ \hat{\varphi}_2^{\mathrm{T}} \\ \vdots \\ \hat{\varphi}_n^{\mathrm{T}} \end{pmatrix}$$

方程组（7-4-12）可表示成矩阵形式

$$\hat{\Gamma}_n \hat{\Phi}_n = \hat{\eta}_n \tag{7-4-14}$$

多变量$AR(n)$模型参数矩阵的估计值为

$$\hat{\Phi}_n = \hat{\Gamma}_n^{-1} \hat{\eta}_n$$

白噪声方差阵的估计是

$$\hat{Q} = \hat{R}(0) - \sum_{i,j=1}^{n} \hat{\varphi}_i \hat{R}(i-j) \hat{\varphi}_i^T = \hat{R}(0) - \hat{\Phi}_n^T \hat{\Gamma}_n \hat{\Phi}_n \tag{7-4-15}$$

上面得到的估计$\hat{\varphi}_1, \hat{\varphi}_2, \cdots\cdots, \hat{\varphi}_n$称为矩估计，也称为耶尔一瓦克尔估计。

(2) 系数矩阵的最小二乘估计。求p变量自回归模型参数的最小二乘估计，即是求一组$\varphi_j, j=1,2,\cdots,n$使得

$$J(\varphi_1, \varphi_2, \cdots, \varphi_n) = \frac{1}{N} \sum_{t=n+1}^{N} [X_t - \sum_{j=1}^{n} \hat{\varphi}_j X_{t-j}]^T$$

$$[X_t - \sum_{j=1}^{n} \hat{\varphi}_j X_{t-j}]$$

达到极小。为此估计量$\hat{\varphi}_j$需满足如下方程：

$$\left. \frac{\partial J}{\partial \hat{\varphi}_j(l,h)} \right|_{\Phi_n = \hat{\Phi}_n} = 0 \qquad \begin{matrix} l,h=1,2,\cdots,p \\ j=1,2,\cdots,n \end{matrix}$$

其中$\varphi_j(l,h)$是系数矩阵φ_j的第l行第h列元素，上式可写成

$$\frac{1}{N} \sum_{t=n+1}^{N} (X_t - \sum_{j=1}^{n} \varphi_j X_{t-j}) X_{t-i}^T = 0 \tag{7-4-16}$$

$$i=1,2,\cdots,n$$

记

$$\hat{R}'(r) = \frac{1}{N} \sum_{t=n+1}^{N} X_t X_{t-r}^T \tag{7-4-17}$$

那么式(7-4-16)为

$$\widehat{R}'(i)=\sum_{j=1}^{n}\widehat{\varphi}_j\widehat{R}'(i-j) \qquad (7\text{-}4\text{-}18)$$

比较式（7-4-17）与（7-4-13），两式虽不同，但当数据个数N充分大时，$\widehat{R}(r)$与$\widehat{R'}(r)$十分接近，因此由式（7-4-18）解出的最小二乘估计与前述矩估计也十分接近。

多变量AR模型参数的矩估计及最小二乘估计都是直接求解形如（7-4-14）的线性方程组，矩阵$\boldsymbol{\Gamma}_n$是$(p\times n)\times(p\times n)$维的，因此计算量很大，尤其当自回归阶数$n$较大时，这种算法几乎难以实现。特别是模型阶数$n$一般并非已知，而是要从1开始逐步增大到某一适当的阶数。对每一步n，都要求出相应的系数矩阵和噪声方差阵，然后根据某种准则确定最佳模型阶数。因此直接求解法的运算量及占用内存都随n增大，在计算机上几乎无法使用。在实际中是采用下述递推算法。

（3）多变量AR模型参数矩阵及白噪声方差的递推算法。为了表明自回归模型参数矩阵$\boldsymbol{\varphi}_j$（$j=1,2,\cdots,n$）随模型阶数而变，我们将式（7-4-10）写成

$$X_t-\boldsymbol{\varphi}_{n1}X_{t-1}-\boldsymbol{\varphi}_{n2}X_{t-2}-\cdots-\boldsymbol{\varphi}_{nn}X_{t-n}=\boldsymbol{\varepsilon}_t \qquad (7\text{-}4\text{-}19)$$

式（7-4-12）的解记为

$$\widehat{\boldsymbol{\Phi}}_{nn}=[\widehat{\boldsymbol{\varphi}}_{n1},\widehat{\boldsymbol{\varphi}}_{n2},\cdots,\widehat{\boldsymbol{\varphi}}_{nn}]^{\mathrm{T}}$$

线性方程组（7-4-14）为

$$\widehat{\boldsymbol{\Gamma}}_n\widehat{\boldsymbol{\Phi}}_{nn}=\widehat{\boldsymbol{\eta}}_n \qquad (7\text{-}4\text{-}20)$$

为了实现递推运算，需引入上式（7-4-20）的对偶方程

$$\widehat{\boldsymbol{\Gamma}}_n\widehat{\boldsymbol{\Psi}}_{nn}=\widehat{\boldsymbol{\zeta}}_n \qquad (7\text{-}4\text{-}21)$$

其中

$$\widehat{\boldsymbol{\zeta}}_n=\begin{bmatrix}\widehat{R}(n)\\ \widehat{R}(n-1)\\ \vdots\\ \widehat{R}(1)\end{bmatrix},\quad \widehat{\boldsymbol{\Psi}}_{nn}=[\hat{\psi}_{n,n-1},\cdots,\hat{\psi}_{n1}]^{\mathrm{T}}$$

多变量AR模型参数矩阵有如下递推公式

$$\hat{\varphi}_{n+1,n+1}=[\hat{R}(n+1)-\sum_{j=1}^{n}\varphi_{nj}\hat{R}(n+1-j)]$$

$$[\hat{R}(0)-\sum_{j=1}^{n}\psi_{nj}\hat{R}(j)]^{-1}$$

$$\psi_{n+1,n+1}=[\hat{R}(n+1)-\sum_{j=1}^{n}\hat{\varphi}_{nj}\hat{R}(n+1-j)]^{T}$$

$$[\hat{R}(0)-\sum_{j=1}^{n}\hat{\varphi}_{nj}\hat{R}(j)^{T}]^{-1} \qquad (7-4-22)$$

$$\hat{\varphi}_{n+1,j}=\hat{\varphi}_{nj}-\hat{\varphi}_{n+1,n+1}\psi_{n,n+1-j} \qquad j=1,2,\cdots,n$$

$$\psi_{n+1,j}=\psi_{nj}-\psi_{n+1,n+1}\hat{\varphi}_{n,n+1-j}$$

初值取为

$$\hat{\varphi}_{11}=\hat{R}(1)\hat{R}(0)^{-1},\quad \psi_{11}=\hat{R}(1)^{T}\hat{R}(0)^{-1} \qquad (7-4-23)$$

模型白噪声方差阵递推公式为

$$\hat{Q}_{n+1}=[I_p-\hat{\varphi}_{n+1,n+1}\psi_{n+1,n+1}]\hat{Q}_n \qquad (7-4-24)$$

$$\hat{Q}_0=\hat{R}(0)$$

其中I_p是p维单位阵。

式（7-4-22）可证明如下：设对某个n已求得$\hat{\Phi}_{nn}$和$\hat{\Psi}_{nn}$，则有

$$\begin{pmatrix}\Gamma_n & \hat{\zeta}_n\\ \hat{\zeta}_n^{T} & \hat{R}(0)\end{pmatrix}\begin{pmatrix}\hat{\Phi}_{n+1,n}\\ \hat{\varphi}_{n+1,n+n}^{+}\end{pmatrix}=\begin{pmatrix}\hat{\eta}_n\\ \hat{R}^{T}_{(n+1)}\end{pmatrix} \qquad (7-4-25)$$

和

$$\begin{pmatrix}\hat{R}(0) & \hat{\eta}_n^{T}\\ \hat{\eta}_n & \Gamma_n\end{pmatrix}\begin{pmatrix}\hat{\Psi}^{T}_{n+1,n+1}\\ \hat{\Psi}_{n+1,n}\end{pmatrix}=\begin{pmatrix}\hat{R}(n+1)\\ \hat{\zeta}_n\end{pmatrix} \qquad (7-4-26)$$

由式（7-4-25）的第一组方程得

$$\hat{\Phi}_{n+1,n}=\hat{\Gamma}_n^{-1}(\hat{\eta}_n-\hat{\zeta}_n\hat{\varphi}_{n+1,n+1}^T)=\hat{\Phi}_{n,n}-\hat{\Psi}_{nn}\hat{\varphi}_{n+1,n+1}^T \tag{7-4-27}$$

因此有式（7-4-22）中的第三式。

由式（7-4-25）的第二组方程得

$$\hat{\zeta}_n^T\hat{\Phi}_{n+1,n}+\hat{R}(0)\hat{\varphi}_{n+1,n+1}^T=\hat{R}^T(n+1)$$

将式（7-4-27）代入得

$$\hat{\zeta}_n^T(\hat{\Phi}_{nn}-\hat{\Psi}_{nn}\hat{\varphi}_{n+1,n+1}^T)+\hat{R}(0)\hat{\varphi}_{n+1,n+1}^T=\hat{R}^T(n+1) \tag{7-4-28}$$

由此可导出

$$\hat{\varphi}_{n+1,n+1}=[\hat{R}(n+1)-\hat{\Phi}_{nn}^T\hat{\zeta}_n][\hat{R}(0)-\hat{\Psi}_{nn}^T\hat{\zeta}_n]^{-1}$$

即式（7-4-22）中的第一式。

由式（7-4-26）的第二组方程得

$$\hat{\psi}_{n+1,n}=\hat{\Gamma}_n^{-1}(\hat{\zeta}_n-\hat{\eta}_n\hat{\psi}_{n+1,n+1}^T)=\hat{\Psi}_{nn}-\hat{\Phi}_{nn}\hat{\psi}_{N+1,n+1}^T \tag{7-4-29}$$

从而有式（7-4-22）中的第四式，利用式（7-4-26）的第一组方程和式（7-4-29）可以导出

$$\hat{\psi}_{n+1,n+1}=[\hat{R}^T(n+1)-\hat{\Psi}_{nn}^T\hat{\eta}_n][\hat{R}(0)-\hat{\Phi}_{nn}^T\hat{\eta}_n]^{-1}$$

注意到 $\hat{\Psi}_{nn}^T\hat{\eta}_n=\hat{\zeta}_n^T\hat{\Phi}_{nn}$不难得出式（7-4-22）中的第二式。

$\hat{Q}$的递推公式（7-4-24）是由式（7-4-15）导出的，因

$$\begin{aligned}\hat{Q}_{n+1}&=\hat{R}(0)-\hat{\Phi}_{n+1,n+1}^T\hat{\Gamma}_{n+1}\hat{\Phi}_{n+1,n+1}\\&=\hat{R}(0)-(\hat{\varphi}_{n+1,1}\cdots,\hat{\varphi}_{n+1,n+1})\hat{\Gamma}_{n+1}\begin{pmatrix}\hat{\varphi}_{n+1,1}^T\\\vdots\\\hat{\varphi}_{n+1,n+1}^T\end{pmatrix}\end{aligned}$$

而由式（7-4-22）第三式知

$$\begin{aligned}(\hat{\varphi}_{n+1,1}\cdots\hat{\varphi}_{n+1,n}\,\hat{\varphi}_{n+1,n+1})=&(\hat{\varphi}_{n,1}\cdots\hat{\varphi}_{nn},0)\\&-\hat{\varphi}_{n+1,n+1}(\hat{\psi}_{nn},\cdots\hat{\psi}_{n,1},-I_p)\end{aligned}$$

代入前式得

$$\hat{Q}_{n+1}=\hat{R}(0)-(\hat{\varphi}_{n1}\cdots\hat{\varphi}_{nn})\Gamma_n\begin{pmatrix}\hat{\varphi}^T_{n1}\\ \vdots\\ \hat{\varphi}^T_{nn}\end{pmatrix}+\hat{\varphi}_{n+1,n+1}$$

$$(\hat{\psi}_{nn}\cdots\hat{\psi}_{n1}-I_p)\Gamma_{n+1}\begin{pmatrix}\hat{\varphi}^T_{n1}\\ \vdots\\ \hat{\varphi}^T_{nn}\\ 0\end{pmatrix}+(\hat{\varphi}_{n1},\cdots\hat{\varphi}_{nn},$$

$$0)\Gamma_{n+1}\begin{pmatrix}\hat{\psi}^T_{nn}\\ \vdots\\ \hat{\psi}^T_{n1}\\ -I_p\end{pmatrix}\hat{\varphi}^T_{n+1,n+1}-\hat{\varphi}_{n+1,n+1}(\hat{\psi}_{nn},$$

$$\cdots\hat{\psi}_{n1},-I_p)\Gamma_{n+1}\begin{pmatrix}\hat{\psi}^T_{nn}\\ \vdots\\ \hat{\psi}^T_{n1}\\ -I_p\end{pmatrix}\cdot\hat{\varphi}^T_{n+1,n+1} \tag{7-4-30}$$

由式（7-4-21)知

$$\Gamma_{n+1}\begin{pmatrix}\hat{\psi}^T_{nn}\\ \vdots\\ \hat{\psi}^T_{n1}\\ -I_p\end{pmatrix}=\begin{pmatrix}\Gamma_n & \hat{\zeta}_n\\ \hat{\zeta}^T_n & R(0)\end{pmatrix}\begin{pmatrix}\hat{\psi}^T_{nn}\\ \vdots\\ \hat{\psi}^T_{n1}\\ -I_p\end{pmatrix}$$

$$=\begin{pmatrix}0\\ \vdots\\ 0\\ \sum_{j=1}^{n}R(j)^T\hat{\psi}^T_{nj}-R(0)\end{pmatrix}$$

因此式（7-4-30）中第三、四项为零，而第五项为

$$-\hat{\varphi}_{n+1,n+1}\left(\hat{R}(0)-\sum_{j=1}^{n}\hat{R}(j)^T\cdot\hat{\psi}^T_{nj}\right)\hat{\varphi}^T_{n+1,n+1}$$

$$=-\hat{\varphi}_{n+1,n+1}\left(R(0)\hat{\varphi}^T_{n+1,n+1}-\sum_{j=1}^{n}R(j)^T\hat{\psi}^T_{nj}\hat{\varphi}^T_{n+1,n+1}\right)$$

$$=-\widehat{\varphi}_{n+1,n+1}[\widehat{R}(n+1)^{T}-\widehat{\zeta}_{n}^{T}(\widehat{\Phi}_{nn}-\psi_{nn}\widehat{\varphi}_{n+1,n+1}^{T})$$

$$-\sum_{j=1}^{n}\widehat{R}(j)^{T}\widehat{\psi}_{nj}^{T}\widehat{\varphi}_{n+1,n+1}^{T}]$$

这里利用了式（7-4-28）。此外，注意到上式右边的

$$\widehat{\zeta}_{n}^{T}(\widehat{\Phi}_{nn}-\widehat{\psi}_{nn}\ \widehat{\varphi}_{n+1,n+1}^{T})=\sum_{j=1}^{n}\widehat{R}(n+1-j)^{T}\widehat{\varphi}_{nj}^{T}$$

$$-\sum_{j=1}^{n}R(j)^{T}\widehat{\psi}_{nj}^{T}\ \widehat{\varphi}_{n+1,n+1}^{T}$$

故式（7-4-30）的第五项为

$$-\widehat{\varphi}_{n+1,n+1}[\widehat{R}(n+1)-\sum_{j=1}^{n}\widehat{\varphi}_{nj}\widehat{R}(n+1-j)]^{T}$$

再利用式（7-4-22）的第二式，便可将其改写成

$$-\widehat{\varphi}_{n+1,n+1}\widehat{\psi}_{n+1,n+1}[\widehat{R}(0)-\sum_{j=1}^{n}\widehat{\varphi}_{nj}\widehat{R}(j)^{T}]$$

$$=-\widehat{\varphi}_{n+1,n+1}\widehat{\psi}_{n+1,n+1}[\widehat{R}(0)-\widehat{\Phi}_{nn}^{T}\widehat{\Gamma}_{n}\widehat{\Phi}_{nn}]$$

将它代入式（7-4-30）的第五项后可得

$$\widehat{Q}_{n+1}=\widehat{R}(0)-\widehat{\Phi}_{nn}^{T}\widehat{\Gamma}_{n}\widehat{\Phi}_{nn}-\widehat{\varphi}_{n+1,n+1}\widehat{\psi}_{n+1,n+1}$$
$$[\widehat{R}(0)-\widehat{\Phi}_{nn}^{T}\widehat{\Gamma}_{n}\widehat{\Phi}_{nn}]$$
$$=(I_{p}-\widehat{\varphi}_{n+1,n+1}\widehat{\psi}_{n+1,n+1})(\widehat{R}(0)-\widehat{\Phi}_{nn}^{T}\widehat{\Gamma}_{n}\widehat{\Phi}_{nn}]$$
$$=(I_{p}-\widehat{\varphi}_{n+1,n+1}\widehat{\psi}_{n+1,n+1})\widehat{Q}_{n}$$

这就是式（7-4-24）

在式（7-4-22）中参数矩阵$\widehat{\psi}_{n+1,j}$，$j=1,2,\cdots,n,n+1$是计算中引入的中间变量，满足对偶方程（7-4-21）。由递推公式（7-4-22）可以看出参数矩阵递推计算步骤如下：首先由多变量数据$\{X_t\}$算出$\widehat{R}(j)$，$0\leqslant j\leqslant n+1$，然后按照下面顺序计算：

$$
\begin{array}{cccccccc}
\hat{\varphi}_{11} & \hat{\varphi}_{22} & \hat{\varphi}_{21} & \hat{\varphi}_{33} & \hat{\varphi}_{31}, \hat{\varphi}_{32} & & \hat{\varphi}_{n+1,n+1} & \hat{\varphi}_{n+1,1}, \cdots, \hat{\varphi}_{n+1,n}, \\
& \rightarrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow \cdots \rightarrow & & \rightarrow \\
\hat{\psi}_{11} & \hat{\psi}_{22} & \hat{\psi}_{21} & \hat{\psi}_{33} & \hat{\psi}_{31}, \hat{\psi}_{32} & & \hat{\psi}_{n+1,n+1} & \hat{\psi}_{n+1,1}, \cdots, \hat{\psi}_{n+1,n}
\end{array}
$$

递推算法与式（7-4-54）直接算法相比，计算量大大地减少了。由式（7-4-14）直接估计$\hat{\varphi}_{nn}$要用到$\hat{\Gamma}_n^{-1}$，而$\hat{\Gamma}_n^{-1}$是$(p\times n)\times(p\times n)$维矩阵，当阶数$n$由低向高增加时，$\hat{\Gamma}_n^{-1}$的阶数及计算量随$n$而增大。但在递推算法式（7-4-22）中需要求逆的只是$p\times p$维矩阵，计算量不随n而增大，而且噪声方差阵也只涉及p维矩阵的运算。由此可见递推算法的计算量确实有极为显著的改进。

3. 多变量自回归模型的预报误差及模型定阶 设p变量$AR(n)$模型为

$$X_t=\varphi_{n1}X_{t-1}+\varphi_{n2}X_{t-2}+\cdots+\varphi_{nn}X_{t-n}+\varepsilon_t \tag{7-4-31}$$

假定已知的观察值为$X_{t-1},X_{t-2},\cdots$，由正交投影可知用真实模型X_t所做的一步线性预报$\hat{X}'_{t-1}(1)$应是

$$\hat{X}'_{t-1}(1)=\varphi_{n1}X_{t-1}+\varphi_{n2}X_{t-2}+\cdots+\varphi_{nn}X_{t-n} \tag{7-4-32}$$

相应的一步预报误差为

$$\tilde{X}'_t=X_t-\hat{X}'_{t-1}(1)=\hat{\varepsilon}_t$$

一步预报误差方差阵为

$$E[\tilde{X}'_t(\tilde{X}'_t)^T]=E[\varepsilon_t\varepsilon_t^T]=Q_n$$

如果模型参数矩阵$\varphi_{n1},\varphi_{n2},\cdots,\varphi_{nn}$用估计值$\hat{\varphi}_{n1},\hat{\varphi}_{n2},\cdots,\hat{\varphi}_{nn}$代替，那么实际一步预报值记作$\hat{X}_{t-1}(1)=\hat{\varphi}_{n1}X_{t-1}+\cdots+\hat{\varphi}_{nn}X_{t-n}$，实际的一步预报误差$\tilde{X}_t$应为

$$\tilde{X}_t=X_t-\hat{X}_{t-1}(1)=\sum_{i=1}^{n}(\varphi_{ni}-\hat{\varphi}_{ni})X_{t-i}+\varepsilon_t$$

为了计算实际的一步预报误差方差阵$D_n=E\tilde{X}_t\tilde{X}_t^T$，我们不加证明引入下面定理[16]。

定理：设$\hat{\Phi}_{nl}$是Φ_{nl}的矩估计或最小二乘估计，则当样本长度充分大时有如下关系

$$E[\sqrt{N}(\Phi_{nl}(i,j)-\hat{\Phi}_{nl}(i,j))]\simeq 0$$

$$E[N(\Phi_{nl_1}(i_1,j_1)-\hat{\Phi}_{nl_1}(i_1,j_1))(\Phi_{nl_2}(i_2,j_2)-\hat{\Phi}_{nl_2}(i_2,j_2))]\cong Q_n(i_1,i_2)\Gamma_n^{-1}(l_1,j_1,l_2,j_2)$$

其中$\Phi_{nl}(i,j)$及$\hat{\Phi}_{nl}(i,j)$分别表示矩阵Φ_{nl}及$\hat{\Phi}_{nl}$的第i行第j列元素。$\Gamma_n(l_1,j_1,l_2,j_2)$表示矩阵$\Gamma_n$的第$(l_1-1)p+j_1$行第$(l_2-1)p+j_2$列的元素，$Q_n(i,j)$及$\Gamma_n^{-1}(l_1,j_1,l_2,j_2)$的含义类似。

我们还需假定$\hat{\Phi}_{nl}(l=1,2,\cdots,n)$与$\{X_t\}$相互独立，即假定我们是用很久以前的数据来估计$\Phi_{nl}$，而用最近获得数据去做预报，由于这两段数据之间相隔时间很长，我们可以近似地把它们看作是独立的。有了这些假定并利用上面定理，我们可以证明一步预报误差方差与白噪声方差有如下近似关系

$$D_n\cong\left(1+\frac{pn}{N}\right)Q_n \qquad (7\text{-}4\text{-}33)$$

由式（7-4-33）可看出，实际模型的一步预报误差比真实模型增大了，这是由于（7-4-33）中包含了用$\hat{\Phi}_{ln}$代替$\Phi_{ln}(l=1,2,\cdots,n)$的误差。用$\hat{Q}_n=\hat{R}(0)-\hat{\Phi}_{nn}^T\hat{\Gamma}_n\hat{\Phi}_{nn}$代替$Q_n$，利用本节定理可以证明

$$E[\hat{Q}_n]\cong\left(1-\frac{pn}{N}\right)Q_n \qquad (7\text{-}4\text{-}34)$$

由式（7-4-34）可知$\hat{Q}_n$是Q_n的渐近有偏估计量，为了克服该估计的偏倚性，我们取

$$\hat{Q}_n=\left(1-\frac{pn}{N}\right)^{-1}(\hat{R}(0)-\hat{\Phi}_{nn}^T\Gamma_n\hat{\Phi}_{nn}) \qquad (7\text{-}4\text{-}35)$$

作为Q_n的估计量，那公式（7-4-35）定义的$\hat{Q}_n$便是Q_n的无偏估计了。

注意

$$\Phi_{nn}^T\Gamma_n\Phi_{nn}=\sum_{i=1}^{n}\sum_{j=1}^{n}\hat{\varphi}_{ni}\hat{R}(i-j)^T\hat{\varphi}_{nj}^T$$

$$=\sum_{i=1}^{n}\hat{\varphi}_{ni}\left(\sum_{j=1}^{n}\hat{R}(i-j)^T\hat{\varphi}_{nj}^T\right)$$

由于$\hat{\varphi}_{nj}$（$j=1,2,\cdots,n$）满足耶尔-瓦克尔方程

$$\hat{R}(i)=\sum_{j=1}^{n}\hat{\varphi}_{nj}\hat{R}(i-j) \quad i=1,2,\cdots,n$$

故有 $$\hat{\Phi}_{nn}^T\hat{\Gamma}_n\hat{\Phi}_{nn}=\sum_{i=1}^{n}\hat{\varphi}_{ni}\hat{R}(i)^T$$

将此式代入（7-4-35），并用$\hat{D}_n$表示用实际估计模型进行预报的一步预报误差，便有

$$\hat{D}_n=\left(1+\frac{pn}{N}\right)\left(1-\frac{pn}{N}\right)^{-1}\left(\hat{R}(0)-\sum_{i=1}^{n}\hat{\varphi}_{ni}\hat{R}(i)^T\right) \qquad (7\text{-}4\text{-}36)$$

前述讨论都是在模型阶数n为已知的条件下进行的，事实上观察序列$\{X_t\}$所适合的AR模型阶数预先并不知道，而是要由观察序列去估计。与单变量AR模型定阶方法相似，对多变量AR模型，我们也是采取FPE准则（最小最终预报误差准则），即以模型的一步预报误差方差来判定模型阶数。我们认为一步预报效果的好坏可以作为检验模型拟合优劣的标准，若一步预报方差阵的行列式越小，则认为模型拟合得越理想。这时的模型阶数，认为

是最适合的阶数。最终预报误差记为

$$FPE_n(X_t)=\det\hat{D}_n=\left(1+\frac{pn}{N}\right)\left(1-\frac{pn}{N}\right)^{-1}$$

$$\det\left[\hat{R}(0)-\sum_{i=1}^{n}\hat{\varphi}_{ni}\hat{R}(i)^T\right]$$

(7-4-37)

我们分别对$n=1,2,\cdots,M$用递推式（7-4-22）求AR模型参数的耶尔一瓦克尔估计。其中M是预先选定的阶数上界，对于长度为N的p变量序列，一般可取为$\frac{N}{10p}\leqslant M\leqslant\frac{N}{5p}$。对得到的模型参数估计值$\hat{\varphi}_{ni}$，$i=1,2,\cdots,n$用式（7-4-37）计算$FPE_n(X_t)$，使最终预报误差达最小的阶数作为$\{X_t\}$模型所适合的阶数。

由式（7-4-37）可以看出，$FPE_n(X_t)$作为阶数n的函数是由两部分构成，其中系数部分$\left(1+\frac{pn}{N}\right)^p\left(1-\frac{pn}{N}\right)^{-p}$ $=\left(\frac{N+np}{N-np}\right)^p$是$n$的单调增加函数。后面的$\det[\hat{R}(0)-\sum_{i=1}^{n}\hat{\varphi}_{ni}\hat{R}(i)^T]=\det\hat{Q}_n$用以表征模型方差的大小。若模型的真实阶数为$n_0$，那么当模型阶数由低到高增加时，$\det\hat{Q}_n$开始时随$n$增大而下降，即$\det\hat{Q}_n\geqslant\det\hat{Q}_{n+1}$，$n=1,2,\cdots,n_0-1$，而当$n>n_0$后，从理论上讲$\det\hat{Q}_n$的值不再减小，因而当$n>n_0$后，$\det\hat{Q}_n$是$n$的单调非增函数。将这两部分结合起来，就会使$FPE_n(X_t)$的曲线在$n_0$处呈现一个凹点。如果序列$\{X_t\}$的真实模型不是$AR$模型，而是用$AR$模型对它作近似描述，那么$\det\hat{Q}_n$就一直随着$n$的增大而减小，然而一般$\det\hat{Q}_n$下降的速率比因子$\left(1+\frac{pn}{N}\right)^p$

$\left(1-\frac{pn}{N}\right)^{-p}$上升的速率来得慢，因此，不难想象把$\mathrm{FPE}_n(X_t)$作为$n$的函数也应存在最低点，使$\mathrm{FPE}_n(X_t)$取最小值相应的$n$，可以确定为是序列$\{X_t\}$所适合的AR模型的阶数，相应的参数矩阵$\hat{\varphi}_{ni}, i=1,2,\cdots,n$就是模型参数的估计值。

为了考察p变量序列$\{X_t\}$中某些分量，比如前l个分量$(l<p)$是否能够描述序列的特征，我们进一步引入下面观察序列部分分量的最终预报误差

$$\mathrm{FPE}_{n,l,p}(X_t)=\left(1+\frac{pn}{N}\right)^l\left(1-\frac{pn}{N}\right)^{-l}$$

$$\det[\hat{R}(0)-\sum_{i=1}^{n}\hat{\varphi}_{ni}\hat{R}(i)^T]_l \qquad (7-4-38)$$

其中$[R(0)-\sum\limits_{i=1}^{n}\hat{\varphi}_{ni}\hat{R}(i)^T]_l$表示$p\times p$维矩阵$\left[\hat{R}(0)-\sum\limits_{i=1}^{n}\hat{\varphi}_{ni}\hat{R}(i)^T\right]$左上角$l$维子矩阵。当$l=r$时，(7-4-38)就是(7-4-37)。如果

$$\min_n \mathrm{FPE}_{n,l,p}(X_t)\geqslant\min_n \mathrm{FPE}_{n,l,l}(X_t)$$

则认为仅考察前l个变元相应分量就够了，没有必要考察p变量序列，换句话说，考察p变量序列与考察l变量序列相比，没有带来明显的效益，也就是说，第$l+1,l+2,\cdots,p$个分量所起作用很小，可以不予考虑。反之，如果

$$\min_n \mathrm{FPE}_{n,l,p}(X_t)\leqslant\min_n \mathrm{FPE}_{n,l,l}(X_t)$$

则表明必须考虑p变量序列，哪个分量都不能忽略。

4. AR模型参数矩阵的计算方法与建模实例 在7.4.2节中已经讲过，通过递推算式（7-4-22）可以计算多变量$AR(n)$模型

的参数矩阵$\hat{\varphi}_{ni}$，$i=1,2,\cdots,n$，但该组算式中涉及到矩阵求逆运算，在计算机上实现时是不方便的。

我们注意到

$$\sum_{j=1}^{n}\hat{\varphi}_{nj}\hat{R}(j)^{T}=\hat{\Phi}_{nn}^{T}\hat{\eta}_{n}=\hat{\eta}_{n}^{T}\hat{\Gamma}_{n}^{-1}\hat{\eta}_{n}$$

这里利用了式（7-4-5），由于上式最右端是对称阵，故左端也是对称阵，类似地

$$\sum_{j=1}^{n}\hat{\psi}_{nj}\hat{R}(j)=\hat{\Psi}_{nn}^{T}\hat{\zeta}_{n}=\hat{\zeta}_{n}^{T}\hat{\Gamma}_{n}^{-1}\hat{\zeta}_{n}$$

这里需用到式（7-4-14）的对偶方程，因此$\sum\limits_{j=1}^{n}\hat{\psi}_{nj}\hat{R}(j)$也是对称阵。将式（7-4-22）两边转置得到

$$\hat{\varphi}_{n+1\ n+1}^{T}=[\hat{R}(0)-\sum_{j=1}^{n}\hat{\psi}_{nj}\hat{R}(j)]^{-1}$$

$$[\hat{R}(n+1)-\sum_{j=1}^{n}\hat{\varphi}_{nj}\hat{R}(n+1-j)]^{T}$$

$$\hat{\psi}_{n+1,\ n+1}^{T}=[\hat{R}(0)-\sum_{j=1}^{n}\hat{\varphi}_{nj}\hat{R}(j)^{T}]^{-1}$$

$$[\hat{R}(n+1)-\sum_{j=1}^{n}\hat{\varphi}_{nj}\hat{R}(n+1-j)]$$

$$j=1,2,\cdots n$$

$$\hat{\varphi}_{n+1,\ j}=\hat{\varphi}_{nj}-\hat{\varphi}_{n+1,\ n+1}\hat{\psi}_{n,\ n+1-j}$$

$$\hat{\psi}_{n+1,\ j}=\hat{\psi}_{nj}-\hat{\psi}_{n+1,\ n+1}\hat{\varphi}_{n,\ n+1-j}$$

$$\hat{\varphi}_{11}^{T}=\hat{R}(0)^{-1}R(1)^{T},\hat{\psi}_{11}^{T}=\hat{R}(0)^{-1}\hat{R}(1) \qquad (7\text{-}4\text{-}39)$$

在递推式（7-4-39）中出现形如$X=A^{-1}B$的运算，其中A是正定对称阵，因此，对正定对称线性方程组$AX=B$求解就可得出X，

所用计算量比矩阵求逆运算小得多。

此外，在用式（7-4-39）进行计算时，我们首先计算

$\hat{\varphi}_{n+1,n+1}$，$\hat{\psi}_{n+1,n+1}$，然后对$j=1,2,\cdots\left[\frac{n}{2}\right]$计算$\hat{\varphi}_{n+1,j}$，$\hat{\psi}_{n+1,n+1-j}$，计算时只用到$\hat{\varphi}_{n,j}$，$\hat{\psi}_{n,n+1-j}$，再算$\hat{\varphi}_{n+1,n+1-j}$，$\hat{\psi}_{n+1,j}$，计算时只用到$\hat{\varphi}_{n,n+1-j}$，$\hat{\psi}_{n,j}$。当$n$为奇数时，还须计算$\hat{\varphi}_{n+1,(n+1)/2}$和$\hat{\psi}_{n+1,(n+1)/2}$，且只用到$\hat{\varphi}_{n,(n+1)/2}$和$\hat{\psi}_{n,(n+1)/2}$。注意到这些特点，在编制程序时就可以做到既节省内存又保证较快的运算速度。

综合上述，多变量自回归模型计算参数矩阵的递推算法与单变量时间序列求偏相关函数的递推算法是完全类似的，所不同的只是多变量情况递推格式更为复杂。

为了便于分析多变量序列相关阵的特性，也为了提高递推算式（7-4-39）的计算精度，我们做如下变换（M是事先选定的自回归最高阶数）

$$\hat{\hat{R}}(i)=\Lambda\hat{R}(i)\Lambda \qquad i=0,1,\cdots,M \tag{7-4-40}$$

其中

$$\Lambda=\begin{pmatrix} \left(\sum_{t=1}^{N}x_{1t}^{2}\right)^{-\frac{1}{2}} & & \bigcirc \\ & \ddots & \\ \bigcirc & & \left(\sum_{t=1}^{N}x_{pt}^{2}\right)^{-\frac{1}{2}} \end{pmatrix}$$

容易验证$\hat{\hat{R}}(0)$的对角线元素为1，而其它元素绝对值都不超过1，同样矩阵$\hat{\hat{R}}(i)$，$i=1,2,\cdots,M$所有元素的绝对值都不超过1，在递推算式（7-4-40）中用$\hat{\hat{R}}(i)$代替$\hat{R}(i)$，并令

$$\hat{\varphi}'_{ni}=\Lambda\hat{\varphi}_{ni}\Lambda^{-1}, \quad \hat{\psi}'_{ni}=\Lambda\hat{\psi}_{ni}\Lambda^{-1} \qquad 1\leqslant i\leqslant n$$

$$\hat{\varphi}'_{n+1,j}=\Lambda\hat{\varphi}_{n+1,j}\Lambda^{-1},\ \hat{\psi}'_{n+1,j}=\Lambda\hat{\psi}_{n+1,j}\Lambda^{-1}\qquad 1\leqslant j\leqslant n+1$$

可以得到关于$\hat{R}(i)$，$\hat{\varphi}'_{ni}$，$\hat{\psi}'_{ni}$，$\hat{\varphi}'_{n+1,j}$，$\hat{\psi}'_{n+1,j}$的递推算式，递推格式与式（7-4-39）完全相同。我们使用这组递推算式进行计算，矩阵$\hat{\hat{R}}(i)$的诸元素绝对值不超过1，因此可以提高计算精度。递推算出的$\varphi'_{n+1,j}$，$\psi'_{n+1,j}$经反变换后得出原模型参数

$$\hat{\varphi}_{n+1,j}=\Lambda^{-1}\hat{\varphi}'_{n+1,j}\Lambda,\ \hat{\psi}_{n+1,j}=\Lambda^{-1}\hat{\psi}'_{n+1,j}\Lambda,\quad 1\leqslant j\leqslant n+1$$

下面给出用气象资料建立多变量AR模型的实例。

我们考察某地区月平均降水、气压、气温、绝对湿度四种观测值，需要分析各气象因素之间的关系。今分别用$\{x_{it}\}i=1,2,3,4$来表示上述气象观察序列，$t=1,2,\cdots,N,N=120$。

表7-2给出了这四种数据资料，表7-3中列出了样本相关阵$\hat{R}(i)$，$i=0,1,\cdots,M$，取$M=20$。表7-3中的每一行是各相关阵（4阶矩阵）的元素按行排列。由表中可以看出，每个序列自身不同时刻之间，各个序列之间相关性都是很强的。

按照式（7-4-38）计算出的最终预报误差列在表7-4中。从该表的第4列可以看出，如果采用四变量模型（降水、气压、气温、绝对湿度）所确定出的自回归模型阶数$n=3$。表中第3、7列中$FPE_{5,3,4}=0.420\times10^{6}$，$FPE_{5,3,3}=0.546\times10^{6}$分别是所在列中最终预报误差的最小值，而前者又比后者小，这说明考察降水、气压、气温三种气象因素之间的相互作用时，绝对湿度的影响是不可忽略的。与此类似，在第2、6、9列中最终预报误差最小值分别是$FPE_{3,2,4}=0.237\times10^{5}$，$FPE_{6,2,3}=0.251\times10^{5}$，$FPE_{6,2,2}=0.268\times10^{5}$，其中第一个是最小的，这说明考虑降水与气压之间的相互关系时，气温与绝对湿度的影响必须考虑。第1、5、8、10列中最终预报误差最小值分别为$FPE_{1,1,4}=0.360\times10^{4}$，$FPE_{3,1,3}=0.346\times10^{4}$$FPE_{2,1,2}=0.351\times10^{4}$，$FPE_{7,1,1}=0.384\times10^{4}$，其中$FPE_{3,1,3}$最小，这说明若只关心雨量的变化规律，绝对湿度的影响可以不予考虑，而气压、气温对雨量的

影响是很显著的。

调用程序MVAR可以计算出使最终预报误差达最小的四变元模型是

$$X_t=\hat{\varphi}_0+\hat{\varphi}_{31}X_{t-1}+\hat{\varphi}_{32}X_{t-2}+\hat{\varphi}_{33}X_{t-3}+\varepsilon_t$$

其中$\hat{\varphi}_{31},\hat{\varphi}_{32},\hat{\varphi}_{33}$由递推算出，$\hat{\varphi}_0=(I-\hat{\varphi}_{31}-\hat{\varphi}_{32}-\hat{\varphi}_{33})\bar{X}_t$，$\bar{X}=(\bar{x}_1,\bar{x}_2,\bar{x}_3)^T$，$\bar{x}_i$是序列$\{x_{it}\}$的均值。具体的估计值如下：

$$\hat{\varphi}_0=(79.654\quad -11.968\quad 38.484\quad 38.852)$$

$$\hat{\varphi}_{31}=\begin{pmatrix} .033 & 2.109 & 6.192 & -.955 \\ .002 & .413 & -.377 & .032 \\ .000 & -.414 & .500 & .004 \\ -.001 & -.433 & .215 & .317 \end{pmatrix}$$

$$\hat{\varphi}_{32}=\begin{pmatrix} -.105 & 2.260 & 1.280 & 1.151 \\ -.009 & .469 & -.210 & .690 \\ .005 & -.423 & -.069 & -.374 \\ .005 & -.511 & .196 & -.764 \end{pmatrix}$$

$$\hat{\varphi}_{33}=\begin{pmatrix} .025 & -3.502 & -4.876 & -2.977 \\ -.004 & .187 & .249 & .322 \\ .001 & -.254 & -.477 & .056 \\ .005 & -.141 & .041 & -.337 \end{pmatrix}$$

表7-2 气象资料的月平均值(1降水,2气压,3气温,4绝对湿度)

1	72.4	59,0	69.9	53.8	132.9	228.3	54.7	37.8	155.5	157.5
2	27.2	22.8	21.2	15.1	10.3	5.9	3.1	5.3	10.6	17.6
3	3.3	5.1	9.5	14.8	19.0	24.1	29.1	28.7	24.9	18.9
4	6.1	7.0	9.4	13.2	17.8	26.5	31.7	31.8	26.5	18.0
1	105.1	16.1	15.4	56.4	13.2	103.4	82.5	100.5	141.7	164.2
2	22.9	24.6	26.8	23.5	19.5	16.1	10.7	5.9	2.4	6.7
3	14.1	6.7	2.0	5.8	8.2	12.5	18.1	23.0	28.0	27.4
4	13.3	7.8	4.9	6.9	7.8	11.6	16.6	23.9	32.4	31.4
1	224.9	45.1	55.9	15.6	.1	26.1	47.2	99.1	195.1	179.4
2	11.6	19.6	23.3	23.7	24.3	26.5	19.6	15.6	9.7	7.2
3	24.0	17.9	11.9	6.5	.5	3.1	9.4	14.4	20.1	22.6
4	25.1	16.5	11.5	7.7	3.4	5.0	8.6	13.6	21.2	23.3
1	43.0	207.6	298.9	30.4	94.5	20.3	81.0	61.8	37.0	170.9
2	4.6	5.6	10.7	20.6	22.0	24.8	24.9	29.7	21.3	14.4
3	27.9	27.9	24.2	16.9	13.0	6.8	5.4	2.0	8.7	16.5
4	30.4	31.5	25.1	14.6	12.6	7.9	7.6	5.8	8.9	16.7
1	115.4	177.8	33.2	37.1	137.8	33.6	5.9	7.8	6.2	109.1
2	12.5	5.4	4.7	5.0	12.1	19.3	23.3	26.1	25.0	23.0
3	18.5	22.6	29.4	27.9	25.8	19.0	11.4	5.9	5.1	5.7
4	17.7	23.0	32.0	30.3	26.7	17.8	9.8	6.9	6.4	7.5
1	39.6	131.7	36.2	175.9	190.4	82.2	12.4	169.4	23.5	53.7
2	21.8	16.6	10.3	7.5	3.7	6.2	12.8	18.7	22.5	26.1
3	7.7	12.1	19.5	22.2	28.1	26.9	22.7	18.0	13.8	4.8
4	7.5	11.7	17.7	22.3	31.2	29.9	21.3	17.0	13.1	6.7
1	23.1	37.9	122.6	106.7	70.3	177.7	222.1	51.3	174.2	4.6
2	25.4	21.6	15.9	14.6	11.8	6.8	5.3	4.6	11.7	18.6
3	4.3	6.5	10.2	13.8	19.0	23.5	27.1	28.6	22.3	18.1
4	6.6	7.9	10.1	13.6	16.7	23.2	29.9	30.8	21.0	15.6
1	55.1	78.8	17.0	44.6	111.0	177.2	159.3	41.6	86.3	.2
2	22.0	25.0	29.0	26.4	19.6	16.8	9.3	4.5	2.7	5.4
3	12.6	6.1	2.0	4.0	8.7	13.6	19.3	24.3	28.0	24.9
4	11.5	7.7	5.4	6.2	9.1	13.3	19.1	25.3	31.0	31.9

续表7-2

1	42.4	10.2	116.8	4.8	43.6	13.3	61.1	99.3	139.5	55.5
2	12.0	19.4	24.6	28.8	24.7	28.3	18.7	18.3	9.4	8.1
3	24.0	18.4	12.5	1.0	2.8	1.8	8.8	13.7	18.7	22.6
4	22.7	15.1	12.1	4.4	5.4	4.7	8.5	11.8	17.9	22.3
1	68.3	83.4	90.9	18.8	47.6	99.6	100.1	80.6	90.0	100.8
2	3.5	5.7	12.8	19.4	22.8	21.0	23.0	24.8	21.2	15.1
3	26.7	27.5	23.7	17.4	13.3	9.5	3.6	2.6	6.8	14.2
4	29.1	29.4	23.6	15.1	12.3	10.6	6.7	8.2	8.3	13.7
1	146.1	55.1	172.6	274.8	125.2	4.8	24.2	9.8	19.4	58.6
2	8.4	6.7	5.2	5.8	11.1	18.6	25.2	27.1	27.2	23.4
3	19.6	22.4	26.7	28.0	24.9	18.0	10.3	4.9	2.0	6.0
4	18.6	21.2	30.1	31.0	26.7	14.9	9.2	6.0	5.3	7.1
1	140.0	38.3	168.0	166.8	104.8	51.9	26.3	17.8	44.5	35.1
2	23.3	17.5	9.8	6.7	3.3	5.7	10.2	19.4	23.7	24.7
3	6.0	13.1	18.7	22.7	27.1	28.1	24.6	18.5	11.8	6.9
4	7.1	11.8	18.1	22.5	30.0	30.6	27.0	17.2	10.8	8.0

表7-3 气象资料的样本相关阵

	R_{11}	R_{12}	R_{13}	R_{14}	R_{21}	R_{22}	R_{23}	R_{24}	R_{31}	R_{32}	R_{33}	R_{34}	R_{41}	R_{42}	R_{43}	R_{44}
R(0)	1.00	-.49	.44	.44	-.49	1.00	-.93	-.94	.44	-.93	1.00	.98	.44	-.94	.98	1.00
R(1)	.24	-.44	.49	.47	-.35	.82	-.93	-.93	.25	-.63	.83	.82	.24	-.63	.82	.82
R(2)	.03	-.35	.42	.41	-.12	.45	-.69	-.67	-.02	-.19	.47	.46	-.02	-.17	.46	.43
R(3)	-.01	-.12	.25	.23	.13	-.03	-.28	-.25	-.29	.30	-.01	-.01	-.28	.32	-.01	-.04
R(4)	-.12	.12	.00	.02	.34	-.47	20	.21	-.45	.69	-.47	-.45	-.44	.68	-.47	-.47
R(5)	-.24	.33	-.25	-.25	.44	-.79	.63	.62	-.47	.90	-.81	-.79	-.47	.87	-.80	-.77
R(6)	-.33	.49	-.43	-.43	.44	-.90	.88	.85	-.39	.88	-.93	-.89	-.40	.84	-.91	-.87
R(7)	-.24	.46	-.46	-.47	.33	-.76	.88	.85	-.23	.61	-.80	-.79	-.23	.59	-.77	-.75
R(8)	-.03	.28	-.38	-.36	.08	-.43	.65	.63	.02	.18	-.45	-.44	.03	.18	-.43	-.43
R(9)	.07	.08	-.23	-.21	-.14	.00	.26	.27	.28	-.29	.00	.01	.28	-.26	.00	-.02
R(10)	.18	-.15	.03	.05	-.33	.43	-.18	-.17	.44	-.66	.44	.44	.43	-.64	.43	.41
R(11)	.15	-.34	.25	.25	-.44	.75	-.58	-.58	.47	-.85	.77	.75	.45	-.84	.74	.74
R(12)	.31	-.44	.41	.39	-.44	.86	-.82	-.82	.37	-.82	.88	.86	.37	-.82	.86	.87
R(13)	.15	-.41	.45	.43	-.32	.73	-.83	-.83	.20	-.57	.75	.73	.21	-.57	.74	.73
R(14)	.01	-.27	.35	.34	-.11	.40	-.62	-.60	-.04	-.16	.42	.40	-.04	-.14	.41	.38
R(15)	-.13	-.03	.15	.15	.16	-.03	-.25	-.22	-.29	.27	-.01	-.01	-.28	.29	-.01	-.04
R(16)	-.19	.15	-.02	-.02	.36	-.42	.19	.19	-.44	.62	-.43	-.41	-.43	.61	-.43	-.42
R(17)	-.20	.33	-.20	-.21	.45	-.70	.56	.55	-.46	.81	-.72	-.70	-.45	.78	-.72	-.69
R(18)	-.26	.39	-.34	-.35	.41	-.80	.78	.75	-.33	.78	-.82	-.81	-.34	.74	-.80	-.77
R(19)	-.20	.39	-.40	-.40	.28	-.67	.77	.75	-.17	.54	-.70	-.70	-.17	.52	-.68	-.65
R(20)	-.15	.24	-.33	-.31	.08	-.38	.57	.55	.04	.16	-.40	-.39	.05	.16	-.38	-.38

表7-4　最终预报误差计算结果

	p=4 l=1	p=4 l=2	p=4 l=3	p=4 l=4	p=3 l=1
0	.420E+04	.201E+06	.195E+07	.612E+07	.420E+04
1	.360E+04	.407E+05	.986E+05	.229E+06	.364E+04
2	.360E+04	.257E+05	.506E+05	.866E+05	.351E+04
3	.363E+04	.237E+05	.423E+05	.606E+05	.346E+04
4	.381E+04	.255E+05	.426E+05	.639E+05	.360E+04
5	.404E+04	.247E+05	.420E+05	.653E+05	.376E+04
6	.414E+04	.257E+05	.465E+05	.654E+05	.382E+04
7	.437E+04	.281E+05	.496E+05	.713E+05	.398E+04
8	.440E+04	.299E+05	.541E+05	.788E+05	.397E+04
9	.428E+04	.299E+05	.575E+05	.818E+05	.397E+04
10	.437E+04	.326E+05	.620E+05	.884E+05	.401E+04
11	.450E+04	.360E+05	.724E+05	.102E+06	.410E+04
12	.464E+04	.395E+05	.843E+05	.113E+06	.417E+04
13	.456E+04	.406E+05	.924E+05	.126E+06	.437E+04
14	.468E+04	.439E+05	.102E+06	.144E+06	.455E+04
15	.501E+04	.487E+05	.117E+06	.167E+06	.474E+04
16	.541E+04	.564E+05	.146E+06	.225E+06	.498E+04
17	.587E+04	.537E+05	.161E+06	.262E+06	.522E+04
18	.639E+04	.756E+05	.190E+06	.328E+06	.541E+04
19	.693E+04	.836E+05	.221E+06	.393E+06	.560E+04
20	.750E+04	.946E+05	.312E+06	.418E+06	.580E+04

	p=3 l=2	p=3 l=3	p=2 l=1	p=2 l=2	p=1 l=1
0	.201E+06	.195E+07	.420E+04	.201E+06	.420E+04
1	.450E+05	.107E+06	.378E+04	.704E+05	.403E+04
2	.327E+05	.753E+05	.351E+04	.407E+05	.409E+04
3	.281E+05	.603E+05	.353E+04	.301E+05	.416E+04
4	.290E+05	.639E+05	.364E+04	.306E+05	.417E+04
5	.252E+05	.546E+05	.374E+04	.276E+05	.408E+04

续表7-4

6	.251E+05	.563E+05	.381E+04	.268E+05	.386E+04
7	.270E+05	.595E+05	.391E+04	.279E+05	.384E+04
8	.277E+05	.631E+05	.395E+04	.274E+05	.390E+04
9	.282E+05	.665E+05	.407E+04	.285E+05	.396E+04
10	.296E+05	.695E+05	.419E+04	.302E+05	.399E+04
11	.306E+05	.729E+05	.427E+04	.319E+05	.406E+04
12	.322E+05	.777E+05	.437E+04	.335E+05	.397E+04
13	.342E+05	.854E+05	.446E+04	.340E+05	.404E+04
14	.369E+05	.935E+05	.455E+04	.347E+05	.411E+04
15	.392E+05	.102E+06	.460E+04	.356E+05	.415E+04
16	.434E+05	.118E+06	.473E+04	.378E+05	.421E+04
17	.475E+05	.133E+06	.490E+04	.405E+05	.427E+04
18	.512E+05	.148E+06	.505E+04	.429E+05	.427E+04
19	.546E+05	.164E+06	.521E+04	.452E+05	.430E+04
20	.572E+05	.180E+06	.529E+04	.473E+05	.416E+04

第八章 一些特定形式的模型

在本章中，对实际中常用的一些具体模型，如趋势性和季节性模型，疏系数模型，门限模型，双线性模型，指数模型，时变AR模型等等，我们将结合实例讨论这些模型的直观背景、物理意义、和在计算机上的实现方法。主要着眼于如何分析使用这些模型，并不偏重于从理论上进行研究探讨。

8.1 趋势性和季节性模型

在许多实际问题中，由量测直接得到的随机序列大多数并不平稳，而是呈现出明显的趋势性或季节性。在前几章我们所讨论的时间序列都来源于均值不变的平稳过程，当然该过程的均值可能是未知的待估参数，但并不是随时间而变化的。在实际观察的序列中可能含有某种随时间稳定增长或衰减的趋势，也可能含有随时间而周期式的变化起伏的趋势。这时，我们就不能认为该序列是均值不变的平稳过程，而需要用更一般的模型来描述，即取

$$x_t=\mu_t+y_t$$

其中μ_t表示x_t中随时间变化的均值，它往往可以用多项式、指数函数、正弦函数等描述，而y_t表示零均值平稳过程，可以用AR-MA模型拟合。做了这样的分解之后，我们就可以用处理平稳过程的方法来分析y_t了。

处理这类问题的方法可分为两大类，一类方法是通过某些处理方法剔除μ_t的变化趋势；另一类方法是具体求出μ_t的拟合形式。如果我们对μ_t的形式并不感兴趣，则可采取第一类方法，

对x_t通过某些运算或处理来剔除μ_t，经处理后，可认为过程$\{x_t\}$只包含平稳的y_t部分，可按平稳过程进行分析和建模，再经反运算可由y_t得出x_t。如果我们不仅要分析y_t，而且也需要知道μ_t的具体形式，那么就要根据观察数据估计μ_t（即求$\widehat{\mu_t}$），然后对残差$\{x_t-\widehat{\mu_t}\}$进行分析，该残差序列可认为是平稳过程，通过对平稳序列的分析处理手段，我们可以得出y_t的估计值$\widehat{y}_t$，于是$\widehat{\mu_t}+\widehat{y_t}$就作为对$x_t$的估计。

关于从观察数据中估计或提取多项式μ_t的方法在1.7节中已做过叙述，本节介绍在时间域内如何剔除μ_t及提取更一般形式μ_t的一些方法。

1 ARIMA模型 在许多实际问题中，我们得到的量测数据序列显然不能近似看作是平稳的，某些序列并不是稳定在同一水平上，而是具有明显的增长或减少趋势。例如图8–1中x_t是日温度的曲线，观察数据序列$\{x_t\}$（见表8–1）有上下起伏的随机变化，但随时间的变化，总的温度变化趋势是增长的。若我们考察日温度的增量$\nabla x_t=x_t-x_{t-1}$，图8–1中的下面一条曲线描述了温度差∇x_t的变化，从图中可以看到，∇x_t可以认为是稳定在零点附近的。

由此不难想象，某些不平稳序列$\{x_t\}$经一阶差分后得到序列$\{\nabla x_t\}$有可能是平稳的，有时可能经若干阶差分后才能平稳化。对x_t作多次差分得到

$$x_t^1=\nabla^d x_t$$

上式中差分算子$\nabla=1-B$，d阶差分算子还可以表示成下面求和形式

$$\Delta^d=(1-B)^d=1-C_d^1B+C_d^2B^2+\cdots+(-1)^{d-1}C_d^{d-1}B^{d-1}+(-1)^dB^d$$

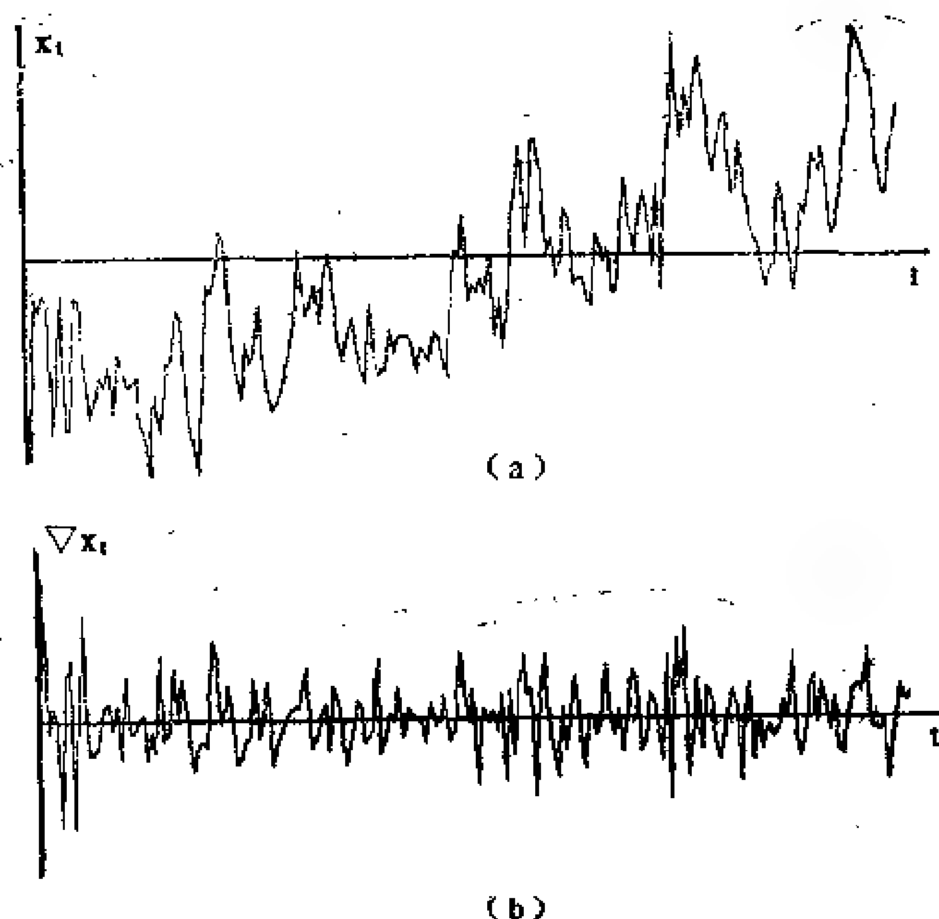


图8—1　日湿度及增量数据曲线

式中C_d^r是从d中取r的组合数，$C_d^r=\dfrac{d!}{r!(d-r)!}$。经过上述差分处理后得到的过程$\{x_t'\}$是平稳的，我们可以对它建立适当的 $ARMA$ 模型

$$\Phi(B)x_t'=\Theta(B)\varepsilon_t$$

其中$\Phi(B)=1+a_1B+\cdots+a_nB^n$，$\Theta(B)=1+b_1B+\cdots+b_mB^m$。对于$x_t$来说则有

$$\Phi(B)(1-B)^d x_t=\Theta(B)\varepsilon_t$$

我们称这种模型为求和自回归滑动平均模型，记作ARIMA(n, d, m)。

表8-1　某地1984年2月1日—8月18日每天中午温度值（℃）

−5.1	−10.1	−.3	−1.1	−.1	−.6	−7.6	−4.6	−.2	−7.3
−7.3	−.1	−.5	−2.8	−5.0	−6.7	−5.8	−4.6	−5.1	−4.0
−6.4	−3.4	−4.3	−5.1	−4.9	−4.6	−7.1	−7.7	−9.7	−5.2
−6.6	−7.4	−3.9	−3.8	−1.1	−1.4	−4.8	−6.4	−8.0	−9.6
−4.2	.2	.0	.8	2.9	2.6	−.4	−3.0	−4.6	−5.6
−2.8	−3.9	−3.4	−.8	−3.8	−5.7	−6.3	−6.0	−5.3	−4.3
−3.4	−1.6	2.0	.1	−1.2	−.8	.3	−1.7	.7	1.9
.4	−2.6	−3.7	−2.4	−1.5	−2.9	−4.2	−4.9	−.7	−3.1
−4.5	−4.1	−4.2	−2.2	−3.3	−2.6	−2.4	−2.4	−2.8	−4.3
−3.2	−3.5	−3.7	−2.3	−2.5	−3.9	−4.7	−.2	2.1	1.8
3.8	−.4	.3	.0	.6	−.1	1.7	−2.4	−.5	−2.9
−.6	3.7	5.4	7.6	2.2	3.6	7.8	7.8	5.8	1.8
2.7	.7	1.4	4.2	3.7	.3	.6	.6	−.6	−.7
2.8	2.7	1.5	2.1	−.2	.0	3.3	5.8	3.4	1.8
3.6	4.9	3.9	1.2	5.5	.1	5.5	7.6	13.6	9.7
7.9	10.5	8.4	10.5	12.2	10.5	7.7	6.2	7.2	9.1
9.2	4.9	5.3	7.6	4.8	4.8	2.8	2.3	1.1	.1
1.2	1.2	5.5	4.3	1.7	−.1	1.6	4.3	3.9	6.1
6.9	5.9	7.3	5.0	2.8	3.1	5.0	7.2	8.8	13.6
13.1	12.3	11.4	11.3	7.2	4.9	5.1	7.1	8.2	9.7

我们简单地说明一下*ARIMA*模型在频率域的含义。若x'_t的谱记作$G_{x_t}(f)$，y_t的谱记作$G_y(f)$，则有

$$G_{x_t}(f)=|1-e^{-j2\pi f}|^{2d}G_y(f)=(2\sin\pi f)^{2d}G_y(f)$$

所以要估计$G_y(f)$可以先求$\hat{G}_{x_t}(f)$，然后计算$\hat{G}_y(f)$，

$$\hat{G}_y(f)=2(\sin\pi f)^{-2d}\hat{G}_{x_t}(f)\qquad f\neq 0。$$

综合上述，对于特定的一类非平稳序列，若该序列是*ARIMA*(n,d,m)序列，则我们可以对它进行d阶差分处理，将其化成平稳*ARMA*序列。对于这类非平稳序列的估计、识别、控制、预报等问题，经差分后化成对相应平稳序列的处理。因此我们将平稳时间序列的分析方法向非平稳序列推进了一步，尽管这并非本质性的推进。

在应用ARIMA模型解决实际问题时，差分阶数d的判断通常要考察原序列和经差分变换后新序列的相关函数。为说明这一

点，在式(8-1-1)两边乘以x_{t-r}并取期望，当$r>m$时有齐次方程

$$\Phi(B)(1-B)^d\rho(r)=0$$

若$\Phi(B)=0$无重根，其解为

$$\rho(r)=A_1+rA_2+\cdots+r^{d-1}A_d+A_{d+1}\ \lambda_{d+1}^r +\cdots+A_{d+n}\ \lambda_{d+n}^r$$

这里$\lambda_{d+1},\cdots,\lambda_{d+n}$的模均小于1。由于$|\rho(r)|$不能超过1，考虑到$r\to\infty$的情况，必须有$A_2=\cdots=A_d=0$，因此$r\to\infty$时$\rho(r)\to A_1=$常数。

一种特定情况就是AR部分的根均为1的模型$(1-B)^d x_t=\varepsilon_t$，其相关函数在任何$r$时$\rho(r)=1$。当实际序列长度有限时，$AR$的根是否准确为1无法判定，然而$AR$根近于1的平稳序列具有"游动"现象，经适当的差分可使其方差显著减小，这种序列实际上是不稳定的。为了说明方便，假定最大的AR根为$\lambda_p=1-\delta$(δ是很小的正数)，当r相当大时

$$\rho(r)\to A_p\lambda_p^r\simeq A_p(1-r\delta)$$

可见当r足够大时不稳定序列的自相关将线性地缓慢衰减（非拖尾）而不是平稳序列那样按指数规律较快地衰减（即不受指数控制）。换言之，样本自相关存在不能很快衰减的倾向时，表示序列是非平稳的。如果从样本偏相关来看，则往往$\phi_{11}(=\rho(1))$近于1。

不妨再看一个ARIMA (0,2,0) 序列，(图8-2画出x_t，∇x_t，$\nabla^2 x_t$，$\nabla^3 x_t$的图形）图8-3a中x_t的自相关由1缓慢衰减，x_t的偏相关表明AR (1) 模型参数a_1近于1。显然需要差分。图8-3b自相关仍然从1线性衰减（虽然衰减较快些），而偏相关再度表示AR(1)模型参数a_1接近1，因此二阶差分可能是必要的。经过二阶差分后得到的自相关和偏相关都符合白噪声过程（见图8-3c），故模型应当是$(1-B)^2x_t=\varepsilon_t$。

如果继续作第三次差分，则样本自相关和偏相关又将偏离白噪声模式，事实上$\nabla^3 x_t = \varepsilon_t - \varepsilon_{t-1}$是一个临界不可逆的MA(1)模型，图8-3d是它的样本自相关和偏相关图形。

本节开头表8-1的温度数据（图8-1）如果经过一次差分，则x_t和∇x_t的自相关和偏相关的变化是很明显的（见图8-4），可取模型为ARIMA（0，1，2）。

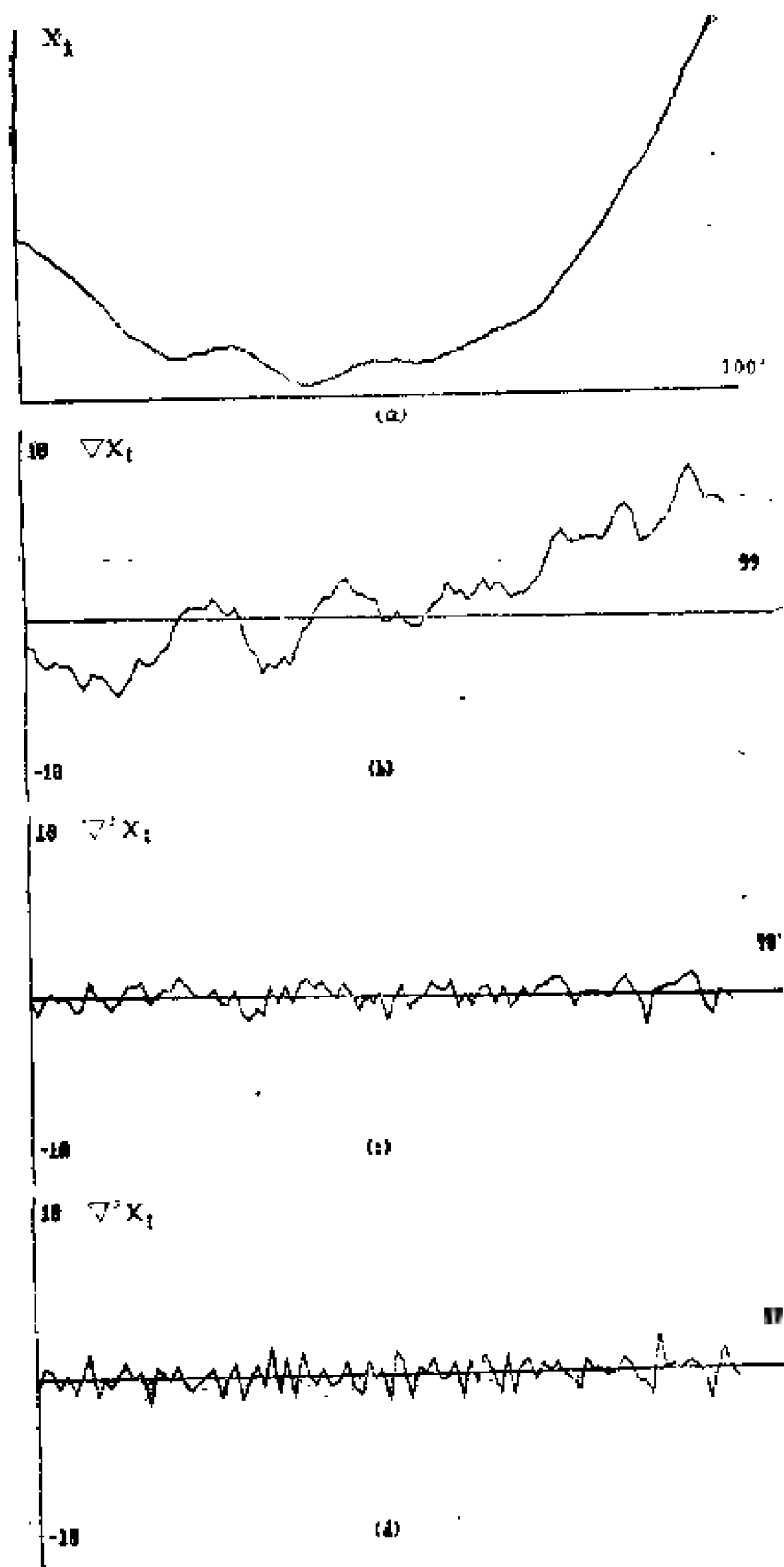


图6-2　ARIMA (0,2,0) 过程的真数据及其差分

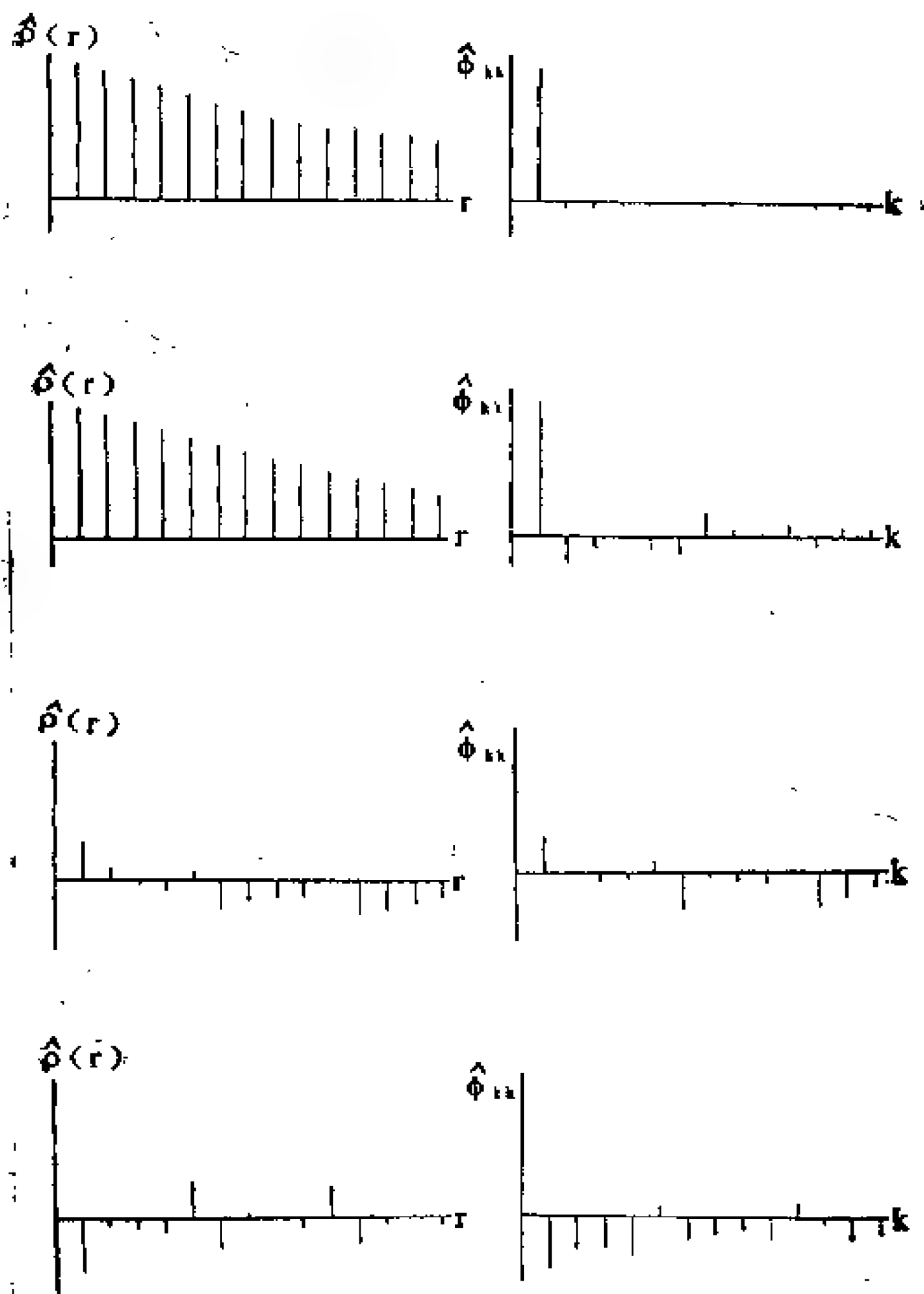
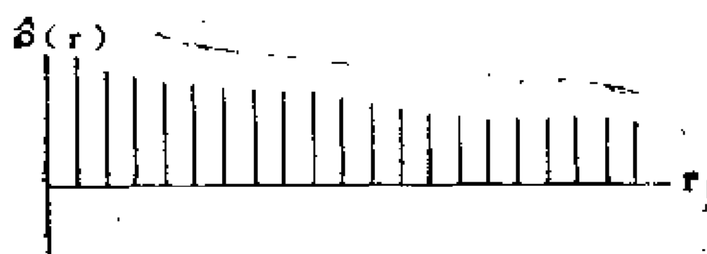
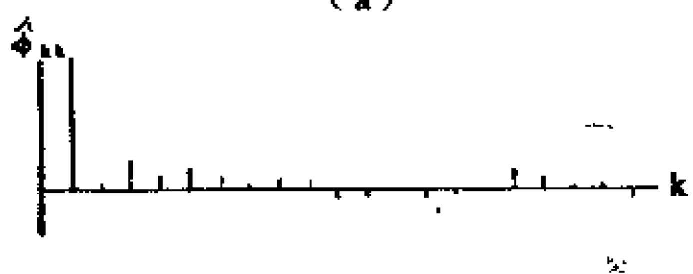


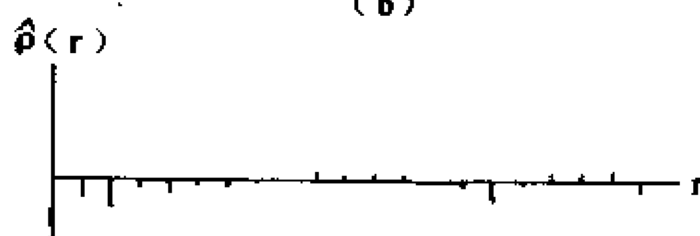
图8-3　仿真数据的自相关和偏相关



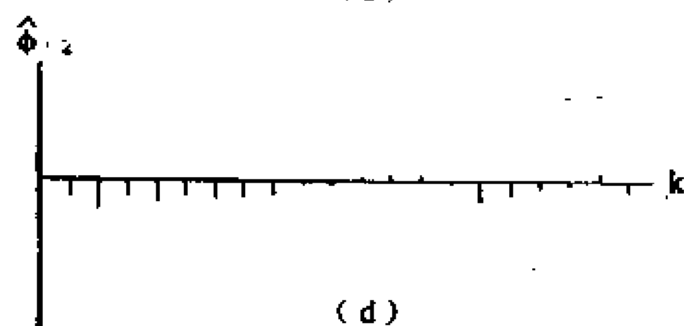
(a)



(b)



(c)



(d)

图8—4　温度数据的自相关、偏相关

2. **乘积型季节性模型** 在本小节中，我们将采用一类疏系数$ARIMA$模型来描述具有周期性变化的随机序列，即所谓乘积型季节性模型。在实际中，我们经常遇到一些依时间而周期性变化的观察序列，例如各种气象观察资料、各种经济信息，各种天文、地震观测数据等等，我们形象地把这类变化规律称为"季节"性变化。在许多问题中，我们往往可以从直观背景及物理变化规律得知数据序列的周期。很自然地想到，如果把某一时刻的观察值与下一周期相应时刻的观察值相减，可能将周期性变化消除掉，例如考察月平均气温的时间序列$\{x_t\}$，显然数据应有年周期(即周期$s=12$月)。若令$x'_t=\nabla_s x_t=(1-B^s)x_t=x_t-x_{t-12}$，则序列$\{x'_t\}$便不再具有明显的$s=12$的周期了。从直观上讲，气温变化是按照一年十二个月周期性地规律变化，但我们把每个月的气温与去年同月气温相减，得到的新序列便是在各月平均气温附近的波动值，因而新序列相当接近于平稳序列。我们把算子∇_s称为季节差分算子，与8.1.1中的差分算子类似，有时季节差分也需要进行若干次。假定原序列经D阶季节差分后得到的序列$\{x'_t\}$有

$$x'_t=\nabla_s^D x_t$$

序列$\{x'_t\}$就周期点上来考察是平稳的，因而可建立关于周期s的时间序列模型

$$U(B^s)x'_t=V(B^s)E_t \qquad (8-1-2)$$

式(8.1.2)称为季节性模型，其中

$$\nabla_s=1-B^s \qquad , \qquad \nabla_s^D=(1-B^s)^D$$

$$U(B^s)=1+u_1B^s+u_2B^{2s}+\cdots+u_pB^{ps}$$

$$V(B^s)=1+v_1B^s+v_2B^{2s}+\cdots+v_qB^{qs}$$

我们还可以考虑稍微广一点的模型，假定式(8-1-2)中的E_t不一定是白噪声，例如是另一个$ARIMA$ (n,d,m)序列，由8.1.1节可知有

$$\Phi(B)\nabla^d E_t = \Theta(B)\varepsilon_t \qquad (8\text{-}1\text{-}3)$$

其中ε_t为白噪声，联合式（8-1-3），（8-1-2）得到

$$\Phi(B)U(B^s)\nabla^d\nabla_s^D x_t = \Theta(B)V(B^s)\varepsilon_t \qquad (8\text{-}1\text{-}4)$$

我们称式（8-1-4）为乘积型季节性模型，其阶数用$(n,d,m)\times(p,D,q)_s$表示。我们可将式(8-1-4)展开成为通常的ARIMA模型形式。例如将

$$(1-B)x_t = (1+b_1B)(1+v_1B^s)\varepsilon_t$$

展开成

$$(1-B)x_t = (1+b_1B+v_1B^s+b_1v_1B^{s+1})\varepsilon_t = \sum_{j=0}^{s+1} b'_j B^j \varepsilon_t$$

这是一个（0，1，$s+1$）阶ARIMA模型，但参数b'_1，b'_2，…，b'_{s+1}有下面约束关系$b'_1=b_1$，$b'_2=\cdots=b'_{s-1}=0$，$b'_s=v_1$，$b'_{s+1}=b_1v_1$。尽管模型的阶数（$s+1$）可能很高，但除b'_1，b'_s，b'_{s+1}以外，其它参数皆是零值，因此是疏系数模型，而且参数b'_{s+1}还满足约束条件$b'_{s+1}=b_1v_1$，实际上只有b'_1和b'_s是自由参数（$b_0=b'_0=1$）。

对于一般形式的（8-1-4），用同样方法也可以化成类似的带约束条件的疏系数ARIMA模型

$$\Phi^*(B)\nabla^d x_t = \Theta^*(B)\varepsilon_t \qquad (8\text{-}1\text{-}5)$$

式（8-1-5）中$\Phi^*(B)$，$\Theta^*(B)$的各参数由相应的$\Phi(B)$，$U(B)$，∇_s^D，$\Theta(B)$，$V(B)$确定，式(8-1-5)相应的模型是疏系数ARIMA（$n+ps+D$，d，$m+qs$）模型，当p，q，D都为零时，便退化成为ARIMA（n，d，m）模型。由此可知8.1.1节中的差分模型是本节乘积型季节模型的特例。

对于乘积型季节性模型进行阶数识别及参数估计，基本上采用波克斯-詹金斯方法，也就是立足于考察随机序列的样本自相关、偏相关函数。这种模型往往由实际物理背景可得知其周期s。如果序列的样本自相关、偏相关函数，二者既不拖尾，也不截

尾，而且不存在衰减趋势，相反地，在相应于周期s整数倍数的点上，自相关（或偏相关）函数出现绝对值相当大的峰值及振荡变化，我们就可以判断数据序列应用乘积型季节模型描述。用$\hat{\sigma}_{x'}(d, D)$记差分阶数d及季节差分阶数D取定后，序列x'_t（$x'_t=\nabla^d\nabla_s^D x_t$）的均方差。我们可选定$d$和$D$的最大变化范围，在此范围内使

$$\hat{\sigma}_{x'}(d_0,D_0)=\min_{d,D}\hat{\sigma}_{x'}(d,D)$$

成立，相应的阶数d_0，D_0便是最佳阶数。或者可采用粗略的办法，考察取不同d、D值时的自相关、偏相关函数，使变换后序列最接近ARMA序列者，被认定是最适宜的差分、季节差分阶数。

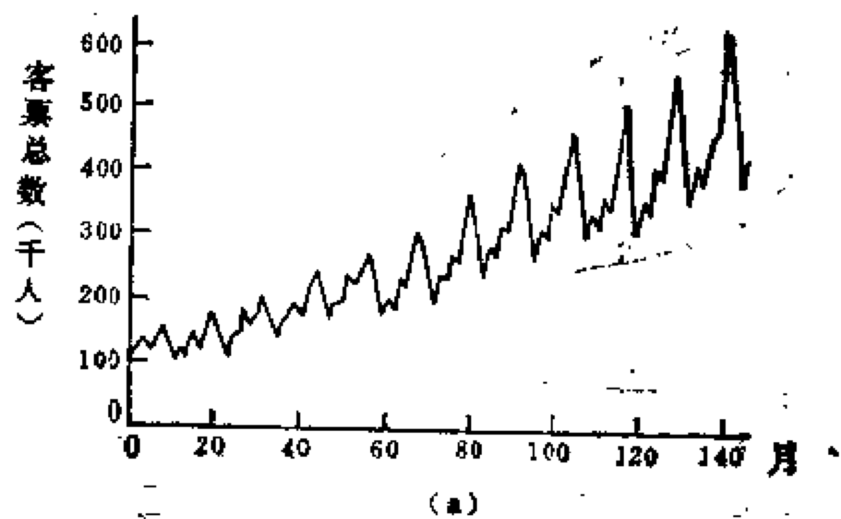
当阶数d、D确定后，我们可以利用第三章中讲过的最小二乘方法，经最优化得到阶数n，m，p,q及$\Phi(B)$，$\Theta(B)$，$U(B)$，$V(B)$中各参数的最佳估计。

下面以1949—1960年每月国际航空公司客票数为例，分析乘积型季节性模型的建立方法。由表8-2及图8-5可以看出，这组数据很明显是周期$s=12$的季节性模型，每年7月、8月旅客人数最多，不同年份的旅客数目有逐年增长的趋势。为了使数据波动幅度基本稳定，将数据取对数（即x_t是原始数据的自然对数），然后分别计算出$\{\nabla x_t\}$，$\{\nabla_{12}x_t\}$，$\{\nabla\nabla_{12}x_t\}$的样本自相关函数（见表8-3和图8-6）。这是因为$\{x_t\}$有线性增长趋势，而$\{\nabla x_t\}$的自相关在迟后为12，24，36，48等处的值都很大，启示我们应做当∇_{12}处理。

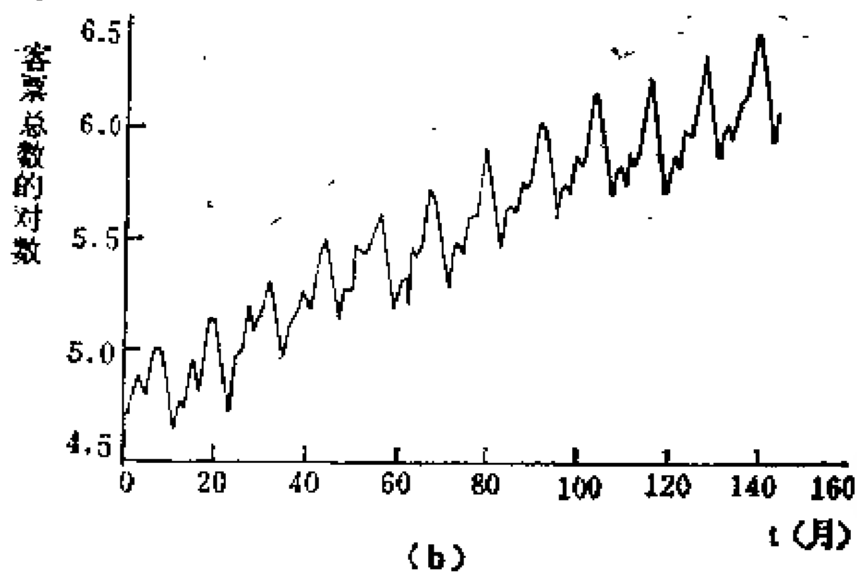
令
$$y_t=\nabla\nabla_{12}x_t$$

为简化问题，我们对新序列$\{y_t\}$进行低阶$ARMA$模型拟合。我们考虑建立如下形式模型

$$y_t=(1+b_1B)(1+v_1B^{12})\varepsilon_t$$



（a）



（b）

图8—5　航空客票数据曲线

a）原数据　　　　b）取对数后的数据

表8.2　1649年—1960国际航空公司客票数（单位：千人）

	1	2	3	4	5	6	7	8	9	10	11	12
1949	112.	118.	132.	129.	121.	135.	148.	148.	136.	119.	104.	118.
1950	115.	126.	141.	135.	125.	149.	170.	170.	158.	133.	114.	140.
1951	145.	150.	178.	163.	172.	178.	199.	199.	184.	162.	146.	166.
1952	171.	180.	193.	181.	183.	218.	230.	242.	209.	191.	172.	194.
1953	196.	196.	236.	235.	229.	243.	264.	272	237.	211.	180.	201.
1954	204.	188.	235.	227.	234.	264.	302.	293.	259.	229.	203.	229.
1955	242.	233.	267.	269.	270.	315.	364.	347.	312.	274.	237.	278.
1956	284.	277.	317.	313.	318.	374.	413.	405.	355.	306.	271	306.
1957	315.	301	356	348.	355.	422.	465.	467.	404.	347.	305	336.
1958	340.	318.	362.	348.	363.	435.	491.	505.	404.	359.	310.	337.
1959	360.	342.	406.	396.	420.	472.	548.	559.	463.	407.	362.	405.
1960	417.	391.	419.	461.	472.	535.	622	606.	508.	461	390.	432.

$$=(1+b_1B+v_1B^{12}+b_1v_1B^{13})\varepsilon_t \qquad (8\text{-}1\text{-}6)$$

即对$\{x_t\}$建立$(0,1,1)\times(0,1,1)_{12}$乘积型季节性模型。利用滑动平均模型的矩估计（见3.2节）可以求得$\hat{b}_1=-0.39$，$\hat{v}_1=-0.48$，若按麦夸特法作最小二乘精估计可得$\hat{b}_1=-0.374$，$\hat{v}_1=-0.565$。

由式（8-1-6）可以推出自协方差函数与模型参数之间的关系如下：

$$\begin{aligned}
R(0)&=(1+b_1^2)(1+v_1^2)\partial_\varepsilon^2\\
R(1)&=b_1(1+v_1^2)\partial_\varepsilon^2\\
R(13)&=R(11)=b_1v_1\partial_\varepsilon^2 \qquad (8\text{-}1\text{-}7)\\
R(12)&=v_1(1+b_1^2)\partial_\varepsilon^2
\end{aligned}$$

其他$R(r)$都等于零。计算表明，按式（8-1 7）求得的$\rho(r)$（$=R(r)/R(0)$）与表8-3中序列$\{\nabla\nabla_{12}x_t\}$的相关函数形状基本符合，因此采用（8-1-6）式模型是可以的。若按模型进行预报并与真值比较，也可以说明模型的适用性。

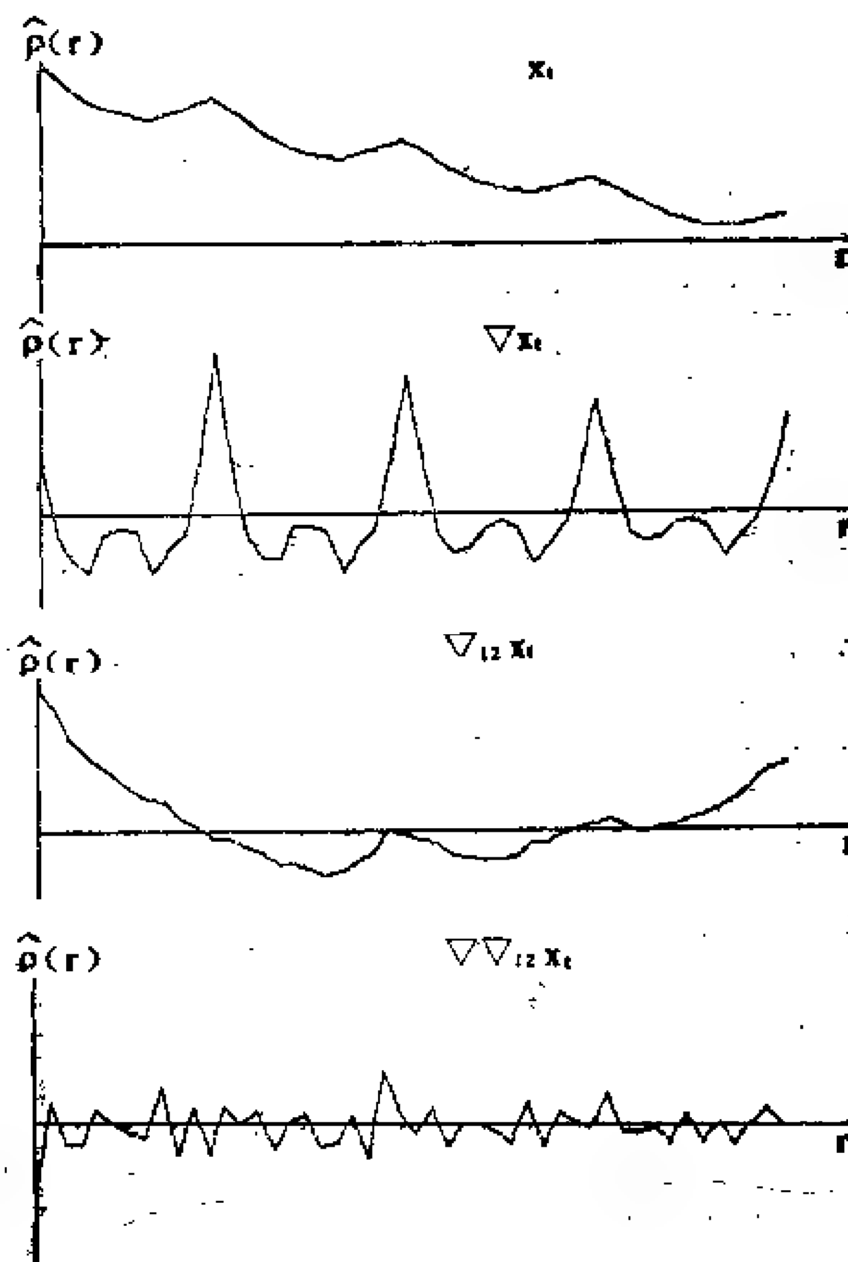

$\hat{\rho}(r)$
x_t
r
$\hat{\rho}(r)$
∇x_t
r
$\hat{\rho}(r)$
$\nabla_{12} x_t$
r
$\hat{\rho}(r)$
$\nabla\nabla_{12} x_t$
r

图8—6　样本自相关曲线

表8-3 $\{x_t\}$, $\{\nabla x_t\}$, $\{\nabla_{12} x_t\}$, $\{\nabla\nabla_{12} x_t\}$的样本自相关

	j	样本自相关函数 $\hat{\rho}(j)$											
(a) X_t	1—12	.948	.876	.807	.763	.714	.682	.663	.656	.671	.703	.743	.760
	13—24	.713	.646	.586	.538	.500	.469	.450	.442	.457	.482	.517	.532
	25—36	.494	.438	.388	.348	.315	.288	.271	.264	.277	.299	.326	.337
	37—48	.303	.254	.211	.172	.138	.111	.091	.082	.089	.104	.124	.133
(b) ∇X_t	1—12	.303	−.102	−.241	−.300	−.094	−.078	−.092	−.295	−.192	−.105	.283	.829
	13—24	.285	−.106	−.222	−.231	−.062	−.066	−.030	−.297	−.163	−.083	.256	.701
	25—36	.257	−.098	−.196	−.174	−.069	−.042	−.078	−.247	−.157	−.047	.195	.580
	37—48	.235	−.106	−.132	−.122	−.053	−.035	−.075	−.227	−.138	−.035	.151	.486
(c) $\nabla_{12} X_t$	1—12	.746	.647	.505	.406	.356	.283	.216	.175	.165	.057	.019	−.044
	13—24	−.035	−.076	−.107	−.176	−.179	−.195	−.235	−.217	−.157	−.117	−.001	−.014
	25—36	−.055	−.068	−.130	−.142	−.155	−150	−.133	−.063	−.063	−.011	.012	.033
	37—48	.061	.008	−.007	.007	.022	.061	.082	.134	.180	.262	.330	.354
(d) $\nabla\nabla_{12} X_t$	1—12	−.310	.095	−.097	−.099	.061	.000	−.056	−.061	.176	−.140	.070	−.134
	13—24	.087	.002	.065	−.109	.000	.044	−.114	−.091	.042	−.157	.258	.053
	25—36	−.051	.081	−102	.002	.000	−.039	−.083	.116	−.099	.057	.010	−.019
	37—48	.155	−.037	−043	−.011	−080	.047	−.072	.011	−.088	.009	.090	−.004

下面再进一步考察序列$\{\nabla\nabla_{12}x_t\}$的相关函数的统计性质。由巴特利公式（见第三章式（3-1-27））得

$$\text{Var}\,\hat{\rho}(r)\cong\frac{1}{N}\{1+2[\rho(1)^2+\rho(11)^2+\rho(12)^2+\rho(13)^2]\}\quad r>13$$

即

$$\sigma=(\text{Var}\,\hat{\rho}(r))^{\frac{1}{2}}\cong 0.11$$

根据正态分布性质，样本落在$\pm\sigma$和$\pm 2\sigma$之间的百分比分别为68.3%和95.5%。我们要检验本例中$\hat{\rho}(r)$，$14\leqslant r\leqslant 48$，这35个样本自相关函数是否服从$N(0, 0.11)$分布，由表8-4可见，这一假设是可以接受的。因此对数据序列进行$\nabla\nabla_{12}$处理是合理的。

表8-4　$\hat{\rho}(r)$的正态性检验

绝对值范围	N（0,0.11）分布的理论频数	样本实际频数
[0,0.11)	23.9	30
[0.11,0.22)	9.5	4
[0.22,+∞)	1.6	1
[0,+∞)	35	35

最后我们来介绍一些实际中常用的分析周期及消除周期分量的方法。

在乘积型季节性模型中，我们引入了差分算子

$$\nabla=1-B$$

和季节差分算子

$$\nabla_s=1-B^s$$

作为算子B的多项式，它们都含有模为1的根。算子∇的根是实数1，算子∇_s的s个根在单位圆上等幅角分布，这是两种特殊情形。

在通常情形下，如果模型相应的算子B的多项式有一对共轭复根，则反映了数据序列含有$\frac{2\pi}{\omega}$的周期，其中ω是共轭复根的正幅角。设

$$1+a_1B+a_2B^2=(1-\lambda_1B)(1-\lambda_2B)$$

$$\lambda_1=\lambda_2=\rho e^{\pm j\omega}$$

利用第二章中$AR(2)$模型根与系数的关系很容易得到λ_1，$\lambda_2=\frac{1}{2}(-a_1\pm\sqrt{a_1^2-4a_2})$，而$a_1=-\lambda_1-\lambda_2=-2\rho\cos\omega$，$a_2=\lambda_1\lambda_2=\rho^2$。表8-5列出了一些典型的周期、算子多项式及特征根。对一些非平稳序列，如建立模型后发现算子多项式含有单位圆上的特征根，则可以通过查表8—5分析出该序列所含周期。如模型的算子多项式含有在单位圆内的复根，我们也可用类似方法找出隐含周期。

表8—5　　几种典型周期、算子多项式及特征根

周期S	算子二次式$1+a_1B+B^2$ ($a_2=1$)	根λ_1、λ_2
2	$1+2B+B^2$	$-1,-1$
3	$1+B+B^2$	$-5+\ i0.866$
4	$1\quad+B^2$	$\pm\ i$
5	$1-0.618B+B^2$	$0.309\pm i0.951$
6	$1-B+B^2$	$0.5\pm i0.866$
7	$1-1.247B+B^2$	$0.624\pm i0.782$
8	$1-1.414B+B^2$	$0.707\pm i0.707$
9	$1-1.532B+B^2$	$0.766\pm i0.643$
10	$1-1.618B+B^2$	$0.809\pm i0.588$

续上表

12	$1-1.732B+B^2$	$0.866\pm j0.500$
14	$1-1.802B+B^2$	$0.901\pm j0.434$
15	$1-1.827B+B^2$	$0.914\pm j0.407$
20	$1-1.902B+B^2$	$0.951\pm j0.309$
21	$1-1.911B+B^2$	$0.956\pm j0.295$
24	$1-1.932B+B^2$	$0.966\pm j0.259$
25	$1-1.937B+B^2$	$0.969\pm j0.249$
28	$1-1.950B+B^2$	$0.975\pm j0.222$
30	$1-1.956B+B^2$	$0.978\pm j0.209$
38	$1-1.968B+B^2$	$0.984\pm j0.178$
50	$1-1.984B+B^2$	$0.992\pm j0.126$
60	$1-1.989B+B^2$	$0.995\pm j0.105$
∞	$1-2B+B^2$	1.00 ，1.00

下面以太阳黑子数为例说明如何从观察数据分析隐含周期.

图8-7及表8-6是1749—1924年逐年的太阳黑子数。图中共有$N=176$个数据，约有$15\frac{1}{2}$个尖峰。对此数据用AR(2)模型拟合，得到自回归参数$a_1=-1.33$，$a_2=0.65$，由式（2-3-41）求出$\cos\omega=-a_1/(2\sqrt{a_2})=0.825$，得出周期约为$2\pi/\omega\simeq 11$年，而$\lambda_1$，$\lambda_2=0.665\pm j0.453$。由于$|a_2|<1$，$|\lambda_1|=|\lambda_2|<1$，因此周期变化带有相当大的阻尼。

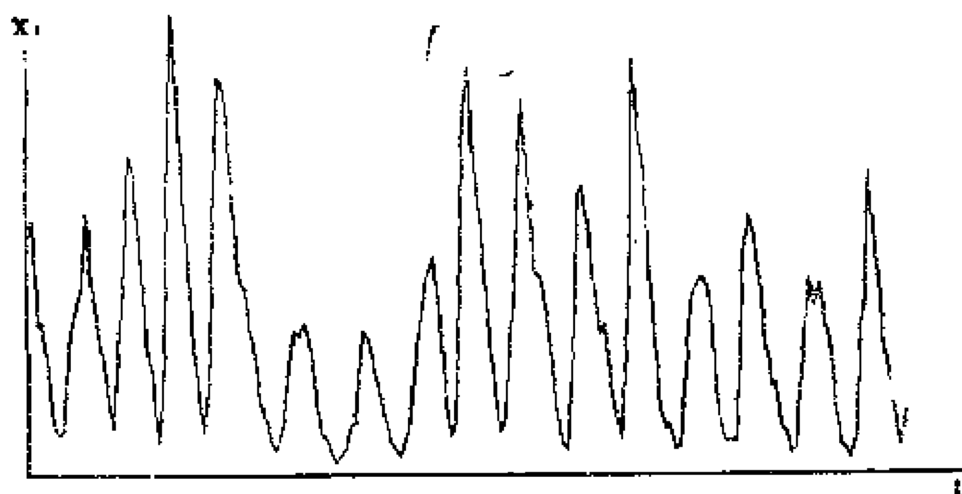


图8—7　太阳黑子数据曲线

表8—6 1749—1924年逐年太阳黑子数

80.9	83.4	47.7	47.8	30.7	12.2	9.6	10.2	32.4	47.6	54.0	62.9
85.9	61.2	45.1	36.4	20.9	11.4	37.8	69.8	106.1	100.8	81.6	66.5
34.8	30.6	7.0	19.8	92.5	154.4	125.9	84.8	68.1	38.5	22.8	10.2
24.1	82.9	132.0	130.9	118.1	89.9	66.6	60.0	46.9	41.0	21.3	16.0
6.4	4.1	6.8	14.5	34.0	45.0	43.1	47.5	42.2	28.1	10.1	8.1
2.5	.0	1.4	5.0	12.2	13.9	35.4	45.8	41.1	30.4	23.9	15.7
6.6	4.0	1.8	8.5	16.6	36.3	49.7	62.5	67.0	71.0	47.8	27.5
8.5	13.2	56.9	121.5	138.3	103.2	85.8	63.2	36.8	24.2	10.7	15.0
40.1	61.5	98.5	124.3	95.9	66.5	64.5	54.2	39.0	20.6	6.7	4.3
22.8	54.8	93.8	95.7	77.2	59.1	44.0	47.0	30.5	16.3	7.3	37.3
73.9	139.1	111.2	101.7	66.3	44.7	17.1	11.3	12.3	3.4	6.0	32.3
54.3	59.7	63.7	63.5	52.2	25.4	13.1	6.8	6.3	7.1	35.6	73.0
84.9	78.0	64.0	41.8	26.1	26.7	12.1	9.5	2.7	5.0	24.4	42.0
63.5	53.8	62.0	48.5	43.9	18.6	5.7	3.6	1.4	9.6	47.4	57.1
103.9	80.6	63.6	37.6	26.1	14.2	5.8	16.7				

若$x_t=\mu_t+y_t$，其中μ_t是以s为周期的周期分量($s>1$)，$\mu_t=\mu_{t-s}$，令$x'_t=(1-B^s)x_t$，则

$$x'_t=(1-B^s)y_t=y_t-y_{t-s}$$

x'_t和y_t的谱之间的关系为

$$G_{x'}(f)=|1-e^{-j2\pi fs}|^2G_y(f)=(2\sin\pi fs)^2G_y(f)$$

由于sin (πfs) 在$f=0$, $1/s$, $2/s$, …处为零，不难看出，经过差分算子 ($1-B^s$) 处理后的x'_t已不含周期为s或其整倍数的分量。因此我们从频谱的角度分析了经差分 ($1-B^s$) 处理可消除$\{x_t\}$中含s及其整倍数的周期分量。实际上，如果我们采用形如 $\left(1-2\cos\left(\frac{2\pi}{p}\right)B+B^2\right)$ 的差分算子，则可剔除$\{x_t\}$中周期为p的分量，而p可以不是采样间隔的整倍数（表8—5是p为整数时的情况）。

还有一种常用的消除周期成份的方法，即所谓"滑动平均滤波"方法。设周期为$2r+1$（r为整数，故周期长度为奇数），显然可以通过（$2r+1$）点的平均来达到消除周期的目的，即

$$x'_t=\frac{1}{2r+1}\left[x_{t-r}+x_{t-r+1}+\cdots+x_t+\cdots+x_{t+r}\right]$$

$$=\frac{1}{2r+1}\left[B^r+B^{r-1}+\cdots+1+B+\cdots+B^{-r}\right]x_t$$

$$=F(B)x_t$$

式中算子式$F(B)$的意义很明显，这样滑动平均的结果只留下常值分量（记作μ），而没有周期分量，此时

$$x'_t=\mu+F(B)y_t$$

由于

$$|F(e^{-i2\pi f})|^2=\frac{\sin^2(2r+1)\pi f}{(2r+1)^2\sin^2\pi f}$$

不难看出，当$f=\frac{n}{2r+1}$（$n=0, 1, 2, \cdots$）时上式为零，因此它确实具有消除周期分量的作用。

3. 组合模型 对于本头开节提到的形式：$x_t=\mu_t+y_t$，通常称μ_t为序列的确定性部分，而称y_t为零均值的平稳随机部分。前面两节所述的$ARIMA$模型和乘积型季节性模型是采取一些办法剔除μ_t部分的作用，而这里要介绍的组合模型则不仅要将趋势性和周期性分量分离出来，而且要给出μ_t的具体表达式。因此最后建立起既有确定性又有随机性部分的组合模型，由这两部分的组合来共同描述某些类型的非平稳过程，常常能得到令人满意的效果。

建立组合模型的方法，简单地说是先用最小二乘法按照某类函数拟合数据序列的确定性部分，从低阶开始，逐渐增加阶数，直到模型无明显改进为止。然后对消除了确定趋势的残量序列建立适宜的$ARMA$（n，m）模型。最后，用前述得到的两部分参数估值作为初值，对确定性部分和$ARMA$部分的所有参数，用非线性最小二乘方法重新估计，得出组合模型的最终估计。

我们对各种组合模型，从简单情形到一般情形分别叙述如下：

（1）线性趋势。数据序列含有线性趋势，即是观察数据在某一直线附近散布，这是组合模型中最简单的情形。我们可以先用熟知的统计方法拟合回归直线，然后对残差建立$ARMA(n,m)$模型，最后把分别估计得到的两部分参数作为初值，用非线性最小二乘方法估计出组合模型的最终参数。

与此类似，对于含有多项式趋势的序列，我们采用适当阶数

的多项式和$ARMA$模型叠加所构成组合模型来拟合。

(2) 指数趋势。数据序列有时呈现指数增长或衰减趋势，这里所讨论的是实指数的情形。我们可用下面组合模型拟合

$$x_t=\sum_{j=1}^{r}A_j e^{K_j t}+y_t \qquad (8\text{-}1\text{-}8)$$

其中K_j，A_j为实数，$\{y_t\}$为ARMA序列。建模的步骤与(1)基本相同。对于$r>1$即多个指数趋势混合的情形，求和阶数r的判别比较麻烦。我们或是根据实际物理背景决定应有若干加项。或是采用使$r=1$，2，3，…由低到高的拟合办法，使r增加到组合模型的残差平方和无显著改进为止。

(3) 周期趋势。如果随机序列不仅含有指数衰减或增长趋势，而且还呈现出某种规律的周期性起伏，则可用如下形式的组合模型去拟合

$$x_t=\sum_{j=1}^{L}R_j e^{\gamma_j t}+\sum_{j=1}^{K}B_j e^{b_j t}\sin(j\omega t+\psi_j)+y_t \quad (8\text{-}1\text{-}9)$$

其中$\{y_t\}$是ARMA (n,m) 序列，L为序列所含有的指数趋势项的项数，K是周期趋势项的项数。ω为基频，由数据的物理性质决定，单位为弧度/秒。B_j和r_j分别表示周期趋势的振幅和相位，$e^{b_j t}$控制周期振荡的增长或衰减趋势。由于在式(8-1-9)中，ψ_j很难用非线性最小二乘法进行估计，因此将模型改成如下形式

$$x_t=\sum_{j=1}^{L}R_j e^{\gamma_j t}+\sum_{j=1}^{K}B_j e^{b_j t}\{C_j\sin(j\omega t)$$

$$+\sqrt{1-C_j^2}\,\mathrm{Cos}(j\omega t)\}+y_t \qquad (8\text{-}1\text{-}10)$$

建模的具体步骤为：首先拟合指数趋势，然后逐个地增添周期趋势项，最后对残差序列$\{y_t\}$建立ARMA模型。各个分模型的参数都估计出来后，再以它们作为初值，对整个组合模型(8-1

-10）统一进行参数估计。在估计过程中需反复使用非线性最小二乘法。由于模型结构比较复杂，待估计参数个数较多，建模过程是比较麻烦的。

（4）组合模型的一般形式。综合前面（1）—（3）的内容可以看出，组合模型可用如下形式统一描述

$$x_t = \sum_{j=0}^{M} A_j e^{K_j t} + y_t \qquad (8\text{-}1\text{-}11)$$

其中$\{y_t\}$是ARMA (n, m) 序列。

如果A_j全部为零，那么

$$x_t = y_t$$

式（8-1-11）成为普通的ARMA（n，m）平稳序列。如果仅有一个A_j非零，且与其相应的K_j等于零，则$\{x_t\}$是均值为常数的平稳序列。如果M个K_j中只存在值很小的实数时，我们将指数函数作泰勒级数展开，则只有前几项较大，因此式（8-1-11）实际上成为具有多项式趋势的模型。当K_j为较大的正负实数时，模型式（8-1-11）成为式（8-1-8），显示出指数增长或衰减的势向。当复指数K_j具有负实部时，模型显示出阻尼的正弦余弦趋向；实部为零时，就是带随机干扰的正余弦趋向；实部为正时，则是振幅不断增大的正余弦趋向。

式（8-1-11）给出的是组合模型的一般形式，在处理具体问题时，我们还是要分析观察数据的特性，并根据实际问题的物理背景，用（1）—（3）中的模型分步进行拟合。

在8.1.2节中，我们对航空公司客票数用波克斯詹金斯方法进行相关分析，建立乘积型季节性模型。现在我们仍以这组数据为例，由简单到复杂，逐步拟合（1），（2），（3）中的各种组合模型，以说明建立组合模型的具体方法。

从图8—5曲线可以看出有理由用带有表示增长趋势的实指数项和表示周期趋势的复指数项的确定性部分的模型来拟合。要拟

合整个确定性加随机的组合模型，我们先从拟合指数增长趋势开始，然后一个一个加上周期趋势项，因此第一个模型取为

$$x_t = R_1 e^{r_1 t} + y_t$$

由数据图8—5量取参数R_1和r_1的初始估计$R_1 \simeq 130$（由$t=0$时$R_1 \simeq x_0$），$r_1 \simeq 0.01$（由$t=1/r_1=100$时$x_t \simeq R_t e^1 = 130$（2.7）

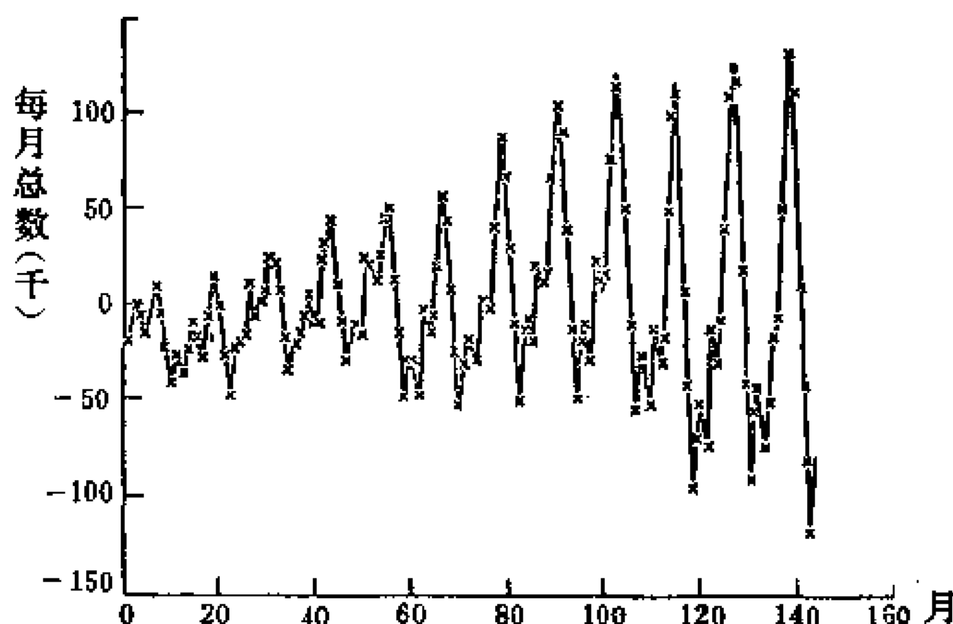


图8—8 消除指数增长趋势后的数据

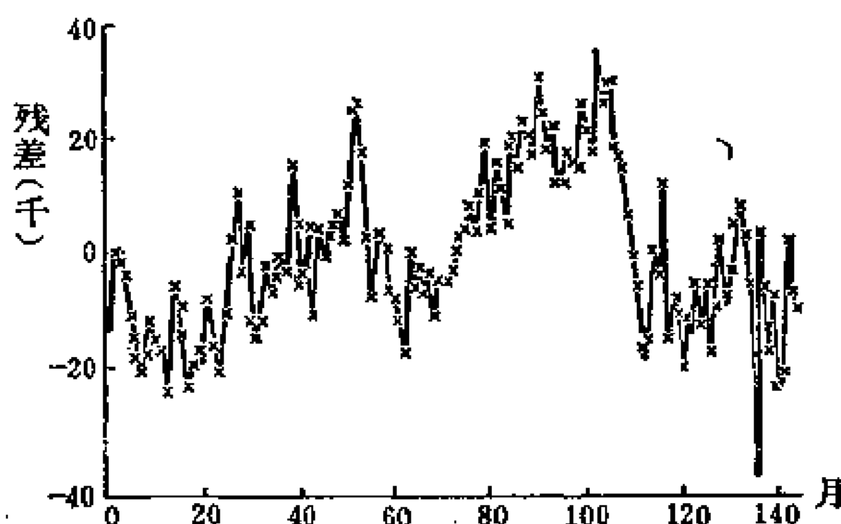


图8—9 消除确定性部分后的残差数据

=351 得)，利用非线性最小二乘估计可得$R_1=130.83\pm10.14$，$r_1=0.0095\pm0.0007$，残差平方和=296250。

第一个模型的y_t如图8－8，可以看出指数增长趋势已被提出，只剩下振幅增长的周期趋势，因此看来用一个实指数就够了（即$L=1$）。下一步是加上一个12个月的周期趋势，即$\omega=2\pi/12$〔月$^{-1}$〕。

$$x_t=R_1e^{r_1t}+B_1e^{b_1t}(C_1\sin\omega t+\sqrt{1-C_1^2}\cos\omega t)+y$$

由图8—9取初始估计$\hat{B}_1=-20$，$\hat{b}_1=0.01$，$\hat{C}_1=0.7$，经非线性最小二乘估计得：

$R_1=131.07\pm5.85$，　　　$r_1=0.0095\pm0.0005$

$B_1=-14.83+6.56$　　　$b_1=0.0136\pm0.0039$

$C_1=0.426\pm0.106$　　　残差平方和=95783

和第一个模型比较起来，残差平方和有显著减小，考虑到这种情况，进一步用$L=1$，$K=2$，3，…时式（8-1-10）模型拟合，直到残差平方和的减小不显著或统计检验认为模型合适为止（详见〔6〕）．由于$K=6$时没有明显改善，故只取$K=5$，即取一年、半年、四个月、一季度和2.4月这五个周期，在用$L=1$，$K=5$的（8-1-10）式模型拟合后的残差已看不出明显的周期趋势（图8-10），其相关函数也表明了它是平稳的（见图8-11）．

随机部分的拟合结果取为$ARMA$（3.3），用分别拟合的结果作为初始估计，然后把确定性和随机部分联合一起进行拟合得参数值如下：

$R_1=127.48\pm0.45$　　　$r_1=0.0096\pm0.0007$

$B_1=-15.89\pm4.063$　　　$b_1=0.0128\pm0.002$

$C_1=0.436\pm0.059$　　　$B_2=-9.503\pm2.4$

$b_2=0.0112\pm0.0022$　　　$C_2=-0.996\pm0.006$

$B_3=3.914\pm1.652$　　　$b_3=0.0113\pm0.0039$

$C_3 = -0.411 \pm 0.122$　　$B_4 = 5.357 \pm 2.642$

$b_4 = -0.006 \pm 0.0472$　　$C_4 = 0.423 \pm 0.425$

$B_5 = 2.383 \pm 1.785$　　$b_5 = 0.0115 \pm 0.0068$

$C_5 = 0.994 \pm 0.025$

$a_1 = -0.087 \pm 0.163$　　$a_2 = -0.14 \pm 0.11$

$a_3 = 0.892 \pm 0.112$　　$b_1 = -0.716 \pm 0.234$

$b_2 = -0.955 \pm 0.227$　　$b_3 = 0.18 \pm 0.256$

残差平方和＝8091

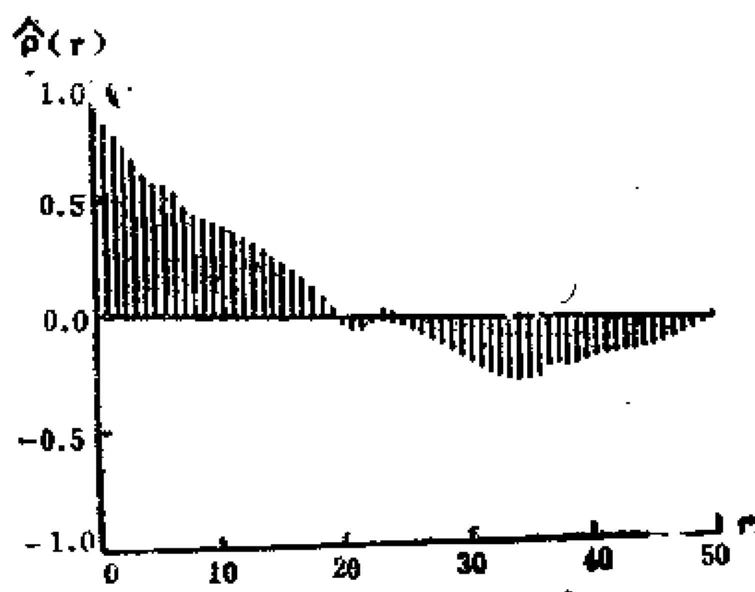


图8—10　图8—9残差数据的自相关

我们对航空公司客票数分别采用了乘积型季节性模型及组合模型进行分析拟合。一般地说，对一组观察数据，我们可以用不同形式的模型去描述。例如，对于具有d阶多项式趋势的序列，我们可以用组合模型去拟合，也可以用$ARIMA$ (n, d, m) 模型拟合；具有周期趋势的序列，也有乘积模型(8-1-4)式和组合模型 (8-1-10) 式两种形式可选择。究竟采用哪种模型描述更适宜，我们要根据数据的物理特性及模型的拟合预报效果来决定。

另外也可根据第三章中讲过的F检验方法或AIC，BIC准则来进行挑选抉择。

最后还需说明一下关于季节性模型的预报问题，利用第六章中对$ARMA$序列的预报方法就可以解决了，但要注意把对$ARMA$序列$\{y_t\}$的预报值还原成原序列$\{x_t\}$的预报值。如采用乘积模型（8-1-4）式则先对$ARMA$序列$\{y_t\}$进行预报，然后利用初值x_1，…，x_D，求出原观察序列$\{x_t\}$的预报值。如果采用组合模型（8-1-10)式，则把确定性部分的预报值与$ARMA$序列$\{y_t\}$的预报值相叠加，便得到观察序列$\{x_t\}$的预报值。

8.2混合回归模型及疏系数模型

本节所介绍的混合回归模型是把经典统计学中的多元回归方法加以推广，引伸来用于时间序列模型的建立，所涉及的数学方法是线性化方法——线性方程组求解，在计算实现上是很方便的。8.2.1节先回顾一般线性回归模型及线性最小二乘方法；8.2.2节介绍常用的线性混合回归模型及其计算，8.2.3节介绍如何将近似AIC准则用于混合回归模型，以及由此产生的疏系数模型。

1．一般线性回归及参数的线性最小二乘估计 设$\{z_t,1\leq t\leq N\}$是因变元的观察序列，$\{x_{it},1\leq t\leq N\}$是第i个自变元（$1\leq i\leq s$）的观察序列。设φ是形式已知而含有未知参数$\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_p)^T$的$s$变元函数，我们用下式表示。

$$z_t=\varphi(x_{1t},\ \cdots,\ x_{st},\ \alpha)+\varepsilon_t \tag{8-2-1}$$

假定ε_1，ε_2，…是独立同分布，且$E\varepsilon_t=0$，$E\varepsilon_t^2=\sigma_\varepsilon^2$，称$\varepsilon_t$为模型残差。式（8-2-1）就是经典统计学中的回归模型，它描述了因变元与自变元之间的依赖关系，若φ是参数$\alpha=(\alpha_1,\cdots,\alpha_p)^T$的线性函数，模型可写成如下形式

$$z_t=\alpha_1\varphi_1(x_{1t},\ \cdots,\ x_{st})+\cdots+\alpha_p\varphi_p(x_{1t},\ \cdots,\ x_{st})+\varepsilon_t \tag{8-2-2}$$

这里$\varphi_1,\cdots,\varphi_p$都是已知的s元函数，且不包含未知数$\alpha_1,\cdots,\alpha_p$。式（8-2-2）就是一般的线性回归模型。注意，这里只需讨论中心化模型（8-2-2），至于非中心化模型

$$z_t=\alpha_0+\alpha_1\varphi_1(x_{1t},\cdots,x_{st})+\cdots+\alpha_p\varphi_p(x_{1t},\cdots,x_{st})+\varepsilon_t \tag{8-2-3}$$

它的常数项参数α_0可由其它参数及观察值计算出。用矩阵表示式（8-2-2），可写成

$$z=A\alpha+\varepsilon \tag{8-2-4}$$

其中

$$A=\begin{bmatrix} \varphi_1(x_{11},x_{12},\cdots,x_{s1}) & \varphi_2(x_{11},x_{21},\cdots,x_{s1})\cdots, & \varphi_p(x_{11},x_{21},\cdots x_{s1}) \\ \varphi_1(x_{12},x_{32},\cdots,x_{s2}) & \varphi_2(x_{12},x_{22},\cdots,x_{s1})\cdots, & \varphi_p(x_{12},x_{22},\cdots x_{s2}) \\ \vdots & \vdots & \vdots \\ \varphi_1(x_{1N},x_{2N},\cdots,x_{sN}) & \varphi_2(x_{1N},x_{2N},\cdots x_{sN})\cdots & \varphi_p(x_{1N},x_{2N},\cdots x_{sN}) \end{bmatrix}=(a_{tj}) \quad \begin{matrix} t=1,2,\cdots,N \\ j=1,2,\cdots,p \end{matrix}$$

$$z=(z_1,\cdots,z_N)^T \quad , \quad \varepsilon=(\varepsilon_1,\cdots,\varepsilon_N)^T$$

我们要求模型式（8-2-4）中参数α的最小二乘估计，即是寻找使下面函数

$$Q(\alpha)=\sum_{t=1}^{N}\left\{z_t-\sum_{j=1}^{p}\alpha_j\varphi_j(x_{1t},\cdots x_{st})\right\}^2 \tag{8-2-5}$$

达到极小值的$\hat{\alpha}$，由线性最小二乘方法知，$\alpha_1,\cdots\alpha_p$应满足正规方程组

$$\alpha_1\sum_{t=1}^{N}a_{t1}a_{tj}+\alpha_2\sum_{t=1}^{N}a_{t2}a_{tj}+\cdots+\alpha_p\sum_{t=1}^{N}a_{tp}a_{tj}=\sum_{t=1}^{N}z_ta_{tj}$$

$$j=1,2,\cdots p \tag{8-2-6}$$

其中$a_{tj}=\varphi_j(x_{1t},\cdots,x_{st})$，写成矩阵形式，正规方程组便是

$$(A^TA)\alpha = A^Tz \qquad (8\text{-}2\text{-}7)$$

若对称矩阵A^TA是正定时，估计值

$$\hat{\alpha} = (A^TA)^{-1}A^Tz$$

可由线性方程组求解得到。

线性方程组（8-2-7)的系数矩阵（向量）是

$$A^TA = \left(\sum_{t=1}^{N} a_{ti}a_{tj}\right)_{1 \le i, j \le p},$$

$$A^TZ = \left(\sum_{t=1}^{N} a_{ti}z_t\right)_{1 \le i \le p}$$

为了提高计算精度，按照通常的办法，应将正规方程中心化、标准化。如考虑有常数项α_0的模型式（8-2-3），则应将正规方程中心化为

$$\begin{cases} s_{11}\alpha_1 + s_{12}\alpha_2 + \cdots + s_{1p}\alpha_p = s_{1z} \\ s_{21}\alpha_1 + s_{22}\alpha_2 + \cdots + s_{2p}\alpha_p = s_{2z} \\ \vdots \\ s_{p1}\alpha_1 + s_{p2}\alpha_2 + \cdots + s_{pp}\alpha_p = s_{pz} \end{cases} \qquad (8\text{-}2\text{-}8)$$

$$\alpha_0 = \overline{z} - \alpha_1\overline{a_1} - \alpha_2\overline{a_2} - \cdots - \overline{a_p}$$

其中

$$s_{ij} = s_{ji} = \sum_{t=1}^{N} a_{ti}a_{tj} - \overline{a_i}\ \overline{a_j}\ , \quad s_{iz} = \sum_{t=1}^{N} a_{ti}z_t - N\overline{a_i}\ \overline{z}$$

$$\overline{a_i} = \left(\sum_{t=1}^{N} a_{ti}\right)/N, \quad \overline{z} = \left(\sum_{t=1}^{N} z_t\right)/N, \quad 1 \le i, \ j \le p$$

进一步需将线性方程组（8-2-8）标准化，得到标准化的正规方程组

$$\begin{cases} s'_{11}\alpha'_1 + s'_{12}\alpha'_2 + \cdots + s'_{1p}\alpha'_p = s'_{1z} \\ s'_{21}\alpha'_1 + s'_{22}\alpha'_2 + \cdots + s'_{2p}\alpha'_p = s'_{2z} \\ \vdots \\ s'_{p1}\alpha'_1 + s'_{p2}\alpha'_2 + \cdots + s'_{pp}\alpha'_p = s'_{pz} \end{cases} \qquad (8\text{-}2\text{-}9)$$

其中

$$s'_{ij}=\frac{s_{ij}}{\sqrt{s_{ii}}\ \sqrt{s_{jj}}},\qquad s'_{iz}=\frac{s_{iz}}{\sqrt{s_{ii}}\ \sqrt{s_{zz}}}\quad 1\leq i,j\leq p$$

$$s_{zz}=\sum_{t=1}^{N}z_t^2-N(\bar{z})^2$$

式（8-2-9）与（8-2-8）中参数α'_i与α_i的换算关系如下：

$$\alpha_i=\alpha'_i\sqrt{\frac{s_{zz}}{s_{ii}}}\qquad 1\leq i\leq p$$

模型的残差平方和可由下式计算

$$Q(\alpha)=s_{zz}(1-\sum_{i=1}^{p}\alpha'_i s'_{iz})$$

综合上述，求解一般线性回归模型参数的计算方法可归纳为如下几个步骤：

①由量测数据$\{z_t,\ x_{ti},\ 1\leq t\leq N,\ 1\leq i\leqslant s\}$各变元之间的相互关系，对任意固定$t$（$1\leqslant t\leq N$）得出式（8-2-4）中$a_{ti}$，$1\leqslant i\leqslant p$及$z_t$的表达式（$1\leqslant t\leqslant N$）。

②列出正规方程组（8-2-9）式

③用平方根法求解正规方程组，得到模型参数α'_1，α'_2，…，α'_p。

④换算得出模型参数α_1，α_2，…，α_p。

由上述可知，处理一般线性回归模型，关键在于步骤①，有了a_{ti}及z_t的表达式，步骤②—④可按照固定程式计算。因此在计算机上实现时，我们可以把①作为一个独立的程序元件模块SUBLSM，而把②—④作为线性回归模型求解的标准程序块（LSME），解决不同的线性回归问题，我们只需相应地置换不同的“元件”。许多实用模型都是借助线性回归模型来实现的，如线

性混合回归模型、门限回归（自回归）模型等等，怎样利用前文提到的程序元件化的方法，我们将在下面各节的具体模型中做进一步说明。

2、 **混合回归模型** 在第三章中我们介绍了序列$\{x_t\}$的自回归模型

$$x_t=\varphi_1 x_{t-1}+\cdots\varphi_n x_{t-n}+\varepsilon_t$$

这种自回归模型与经典统计学中的回归模型不同，它描述了因变元本身"当前"与"过去"的统计依赖关系。结合8.2节中的一般回归模型，我们可以看到，一般回归模型与自回归模型都有各自的局限性；一般回归模型没有体现出因变元自身前后及因变元与各自变元"过去"的依赖关系，而自回归模型则没有考虑其它变元对因变元的作用。在7.7.2节中讨论的多变量自回归模型，可在一定程度上弥补前面两种模型的欠缺，但由于多变量自回归模型的具体形式及递推算法的限制，必须按照一种固定的程式计算，可能待估计的模型参数过多，因而不够理想。

本节介绍的混合回归模型能够弥补单纯回归模型与自回归模型的不足，而且由具体计算方法可知，它比多变量自回归模型构造灵活，可以用较少的参数建立阶数较高的模型。

在实际问题中，我们经常遇到相互影响的多个观察序列，如考察某地的逐月降水量，它不仅与该月的温度、气压、湿度有关，而且还与温度、气压、湿度的历史值及雨量自身历史值有关。对于这类观测资料就应考虑采用混合回归模型拟合。本节主要考虑线性混合回归模型，因为许多混合回归模型可以用线性混合回归近似，这样我们可以充分利用8.2.1节中所讲的线性回归模型这一方便而得力的工具。

下面介绍几种常用的混合回归模型，并考虑如何用8.2节中的程序元件化方法实现参数估计。

(1) 混合线性回归模型。 综合线性回归与自回归模型的特

点，我们定义线性混合回归模型为如下形式

$$z_t=\alpha_0+\alpha_1^{(0)}z_{t-1}+\cdots+\alpha_{p_0}^{(0)}z_{t-p_0}+\alpha_1^{(1)}x_{1,t-1}+\cdots$$
$$+\alpha_{p_1}^{(1)}x_{1,t-p_0}+\cdots+\alpha_1^{(s)}x_{s,t-1}+\cdots+\alpha_{p_s}^{(s)}x_{s,t-p_s}$$
$$+\varepsilon_t \qquad (8\text{-}2\text{-}10)$$

由模型（8-2-10）式可知，因变元z_t不仅依赖于其自身历史值z_{t-1}，…，z_{t-p_0}，而且还依赖于自变元的某些历史值$x_{1,t-1}$，…，$x_{1,t-p_1}$，…，$x_{s,t-1}$，…，$x_{s,t-p_s}$

为了程序编排方便，我们用$\{x_{i,t},1\leqslant t\leqslant N,\ 1\leqslant i\leqslant s\}$来表示观察序列（包括因变元和自变元），这时变量的个数应变成$s+1$，但为了简便起见，我们仍记做s，并认定第j个变元作为因变元，于是模型（8-2-10）式改写成

$$x_{jt}=\alpha_0+\alpha_1^{(1)}x_{1,t-1}+\cdots+\alpha_{p_1}^{(1)}x_{1,t-p_1}+\alpha_1^{(2)}x_{2,t-1}$$
$$+\cdots+\alpha_{p_2}^{(2)}x_{2,t-p_2}+\cdots\alpha_1^{(s)}x_{s,t-1}+\cdots\alpha_{p_s}^{(s)}x_{s,t-p_s}+\varepsilon_t$$
$$(8\text{-}2\text{-}11)$$

令$m=p_1+p_2+\cdots+p_s$，$p=\max(p_1,\ p_2,\ \cdots,\ p_s)$，取$p+1\leqslant t\leqslant N$以保证式（8-2-11）中各变量下标为正，于是线性回归模型（8-2-4）式中相应的量为

$$a_{tl}\Leftarrow x_{l,t-h}$$
$$z_t\Leftarrow x_{jt}$$

其中$j=\sum_{i=1}^{l-1}p_i+h\quad 1\leqslant h\leqslant p_l$，$1\leqslant l\leqslant s$，并规定求和的上限小于下限时求和式为零。我们很容易编制出混合线性回归模型（8-2-1）式相应的程序元件（SUBLSM），并调用线性回归子程序（LSME），立刻可以得出混合线性回归模型参数的最小二乘估计。

待估计参数个数为m个。当量测变量个数较多且考虑延迟步数较长时，参数个数是相当多的。另外，各个观测序列取不同的迟后，就可以产生出许多不同的模型。譬如我们取定最大迟后为

p，那么每个变量的迟后p_j都有：$0\leqslant p_j\leqslant p$，$j=1, 2, \cdots, s$。因此，我们可以组合成包含不同迟后量的许多混合线性回归模型，所有可能的模型总个数是2^{ps}个。如果我们对这样多的模型进行拟合，并用AIC准则挑选最佳线性混合回归模型，计算量是很大的。在本章下节中我们将介绍解决这一问题的近似方法。

(2)多变量自回归模型。假定多变量观察序列$\{x_{jt}, 1\leqslant t\leqslant N, 1\leqslant j\leqslant s\}$的每一个变量都适合于式（8-2-11）给出的线性混合回归模型

$$
\begin{aligned}
x_{jt}=\alpha_j^{(0)}&+\alpha_{j1}^{(1)}x_{1,t-1}+\alpha_{j1}^{(2)}x_{1,t-2}+\cdots+\alpha_{j1}^{(p)}x_{1,t-p}+\\
&+\alpha_{j2}^{(1)}x_{2,t-1}+\alpha_{j2}^{(2)}x_{2,t-2}+\cdots+\alpha_{j2}^{(p)}x_{2,t-p}+\\
&+\cdots+\alpha_{js}^{(1)}x_{s,t-1}+\alpha_{js}^{(2)}x_{s,t-2}++\alpha_{js}^{(p)}x_{s,t-p}\\
&+\varepsilon_t^{(j)} \qquad 1\leqslant j\leqslant s \qquad (8\text{-}2\text{-}12)
\end{aligned}
$$

最大迟后取为p。式（8-2-12）的矢量形式为

$$
\boldsymbol{x}_t=\boldsymbol{a}_0+\boldsymbol{A}_1\boldsymbol{x}_{t-1}+\boldsymbol{A}_2\boldsymbol{x}_{t-2}+\cdots+\boldsymbol{A}_p\boldsymbol{x}_{t-p}+\boldsymbol{\varepsilon}_t \quad (8\text{-}2\text{-}13)
$$

其中

$$
\boldsymbol{a}_0=(\alpha_1^{(0)},\cdots,\alpha_s^{(0)})^T,\quad \boldsymbol{x}_{t-i}=(x_{1,t-i},x_{2,t-i},\cdots,x_{s,t-i})^T \quad 1\leqslant i\leqslant p
$$

$$
\boldsymbol{A}_1=\begin{pmatrix}\alpha_{11}^{(1)} & \alpha_{12}^{(1)}\cdots\alpha_{1s}^{(1)}\\ \alpha_{21}^{(1)} & \alpha_{22}^{(1)}\cdots\alpha_{2s}^{(1)}\\ \vdots & \\ \alpha_{s1}^{(1)} & \alpha_{s2}^{(1)}\cdots\alpha_{ss}^{(1)}\end{pmatrix},\qquad
\boldsymbol{A}_2=\begin{pmatrix}\alpha_{11}^{(2)} & \alpha_{12}^{(2)}\cdots\alpha_{1s}^{(2)}\\ \alpha_{21}^{(2)} & \alpha_{22}^{(2)}\cdots\alpha_{2s}^{(2)}\\ \vdots & \\ \alpha_{s1}^{(2)} & \alpha_{s2}^{(2)}\cdots\alpha_{ss}^{(2)}\end{pmatrix},
$$

$$
\boldsymbol{A}_p=\begin{pmatrix}\alpha_{11}^{(p)} & \alpha_{12}^{(p)}\cdots\alpha_{1s}^{(p)}\\ \alpha_{21}^{(p)} & \alpha_{22}^{(p)}\cdots\alpha_{2s}^{(p)}\\ \vdots & \\ \alpha_{s1}^{(p)} & \alpha_{s2}^{(p)}\cdots\alpha_{ss}^{(p)}\end{pmatrix},\qquad
\boldsymbol{\varepsilon}_t=(\varepsilon_t^{(1)},\varepsilon_t^{(2)},\cdots,\varepsilon_t^{(s)})^T
$$

由式（8-2-13）可知，矢量形式模型是多变量AR（p）模型（见第七章7.4节），但每个分量所适合的模型（8-2-12）式都是线性混合回归模型。如果我们直接用多变量$AR(p)$模型进行拟合，$\boldsymbol{A}_1$，…，$\boldsymbol{A}_p$所包含的模型参数过多，特别对观察样本不太长的

情形，势必使参数估计误差增大。即便按每个分量用普通的线性混合回归模型（8-2-12）式拟合，计算量依然很大，而且参数个数仍很多。一种容易实现且节省存贮和计算量的做法是：既把多变量自回归模型分解成（8-2-12）式，又对每一分量用8.3节将介绍的疏系数线性混合回归模型拟合（后面将介绍挑选变元的方法和近似AIC准则），然后将各分量相应的模型组合起来，便得到疏系数多变量自回归模型。

（3）增广混合回归模型。设$\{z_t\}$是因变元观察序列，$\{x_{jt}\}$，$1\leqslant j\leqslant s$是自变元观察序列。我们定义下式为增广混合回归模型

$$z_t=\alpha_1 x_{1t}+\alpha_2 x_{2t}+\cdots+\alpha_p x_{pt}+\eta_t \qquad (8\text{-}2\text{-}14)$$

$$\eta_t=\beta_1\eta_{t-1}+\cdots+\beta_n\eta_{t-n}+\varepsilon_t \qquad (8\text{-}2\text{-}15)$$

其中$\{\varepsilon_t\}$是相互独立的正态白噪声序列，若$\beta_1,\cdots,\beta_n$满足平稳性条件，则式（8-2-15）为n阶自回归模型，称式（8-2-14）为平稳增广混合回归模型。当$\eta_t=\varepsilon_t$（$t=1,2,\cdots,N$）时，式(8-2-14)退化成一般的线性回归模型。与此同时，当x_{1t}，$\cdots x_{pt}$取某些特定形式时就成为线性混合回归模型或多变量自回归模型。

由上述可知，增广混合回归模型是一种代表形式广泛的线性回归模型。这种模型用线性最小二乘方法是无法解决的，因为对于参数$\alpha=[\alpha_1,\cdots\alpha_p]^T$，$\beta=[\beta_1,\cdots\beta_n]^T$模型（8-2-14）的拟合残差平方和为

$$\begin{aligned}Q(\alpha,\beta)=\sum_{t=n+1}^{N}\{&(z_t-\alpha_1x_{1t}-\cdots-\alpha_px_{pt})-\beta_1(z_{t-1}\\&-\alpha_1x_{1,t-1}-\cdots-\alpha_px_{p,t-1})-\beta_2(z_{t-2}\\&-\alpha_1x_{1,t-2}-\cdots-\alpha_px_{p,t-2})-\cdots\\&-\beta_n(z_{t-n}-\alpha_1x_{1,t-n}-\cdots-\alpha_px_{p,t-n})\}^2\end{aligned}$$

极小化残差平方和求参数估计值，得到的

$$\frac{\partial}{\partial \alpha_j} Q(\boldsymbol{\alpha}, \boldsymbol{\beta}) = 0 \qquad j = 1, 2, \cdots, p$$

$$\frac{\partial}{\partial \beta_j} Q(\boldsymbol{\alpha}, \boldsymbol{\beta}) = 0 \qquad j = 1, 2, \cdots n \tag{8-2-16}$$

是关于α_j，$1 \leqslant j \leqslant p$，和$\beta_j$，$1 \leqslant j \leqslant n$的非线性方程，因此这是非线性最小二乘求解问题，当然我们可以用迭代方法求数值解，但是比较方便易行的办法是用线性求解方法做近似拟合，具体做法是：

①先拟合回归模型（8-2-14），把式（8-2-14）中的η_t看作式（8-2-4）中的ε_t，给出参数$\boldsymbol{\alpha}$的估计$\widehat{\boldsymbol{\alpha}}$。具体实现时只须在线性回归模型（8-2-4）中令

$$a_{ti} \Leftarrow x_{it}, \qquad z_t \Leftarrow z_t \qquad 1 \leqslant i \leqslant p,\ 1 \leqslant t \leqslant N$$

编制子程序SUBLSM，然后调用程序LSM立刻可得出估计值$\widehat{\boldsymbol{\alpha}}$。如果需要的话，也可以对模型阶数$p$及变元进行优选。

②利用①中的拟合值$\widehat{\boldsymbol{\alpha}}$求出拟合残差

$$\widehat{\eta_t} = z_t - \widehat{\alpha}_1 x_{1t} - \widehat{\alpha}_2 x_{2t} - \cdots - \widehat{\alpha}_p x_{pt} \quad t = 1, 2, \cdots, N$$

对残差$\widehat{\eta_1}$，…，$\widehat{\eta_N}$进行自回归拟合，得出式（8-2-15）模型阶数及参数的估计值。

用上述方法给出的估计值$\widehat{\boldsymbol{\alpha}}$，$\widehat{\boldsymbol{\beta}}$在最小二乘意义下，不如模型（8-2-14），（8-2-15）的严格解好，但是这种近似方法比严格最小二乘容易实现。事实上，在一定的条件下当样本个数N趋于无穷时，近似估计$\widehat{\boldsymbol{\alpha}}$，$\widehat{\boldsymbol{\beta}}$收敛于模型参数的真值。

3. 挑选回归变元的近似最小信息量准则及疏系数模型 考虑形如式（8-2-4）的线性回归模型，我们将它表示成如下形式

$$z_t = \alpha_1 x_{1t} + \alpha_2 x_{2t} + \cdots + \alpha_n x_{nt} + e_t \tag{8-2-17}$$

其中$\{z_t\}$是因变元观察序列，$\{x_{it}\}$，$1\leqslant i\leqslant n$可以是直接量测到的自变元，也可以是某些自变元按照已知函数关系生成的变量，我们称之为导出变元，利用式（8-2-9）便可以得到和模型(8-2-17）式相应的正规方程。为了简化记号，我们将式（8-2-9）中各变量的上标“1”省去，并令$y_i=s_{iz}$，$1\leqslant i\leqslant n$，于是得到正规方程系数的加边矩阵

$$S=\begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1n} & y_1 \\ s_{21} & s_{22} & \cdots & s_{2n} & y_2 \\ \vdots & & & & \vdots \\ s_{n1} & s_{n2} & \cdots & s_{nn} & y_n \\ y_1 & y_2 & \cdots & y_n & y_0 \end{pmatrix} \tag{8-2-18}$$

此处取$y_0=1$。例如考虑式（8-2-17）的某一特定模型

$$z_t=\alpha_1 x_{1t}+\alpha_4 x_{4t}+\varepsilon_t$$

与该模型相应的加边矩阵是

$$S^{(1)}=\begin{pmatrix} s_{11} & s_{14} & y_1 \\ s_{41} & s_{44} & y_4 \\ y_1 & y_4 & y_0 \end{pmatrix} \tag{8-2-19}$$

实际上，若先从矩阵（8-2-18）中取第1、第4、第$n+1$列，得到下面矩阵

$$\begin{pmatrix} s_{11} & s_{14} & y_1 \\ \vdots & \vdots & \vdots \\ s_{41} & s_{44} & y_4 \\ \vdots & \vdots & \vdots \\ s_{n1} & s_{n4} & y_n \\ y_1 & y_4 & y_0 \end{pmatrix}$$

再从这个矩阵中取第1、4、$n+1$行便得到式(8-2-19)矩阵$S^{(1)}$。由这个例子可以看出，我们在变元$x_{1t},x_{2t},\cdots,x_{nt}$中挑选其中部分变元建立回归模型(称为模型(8-2-17）式的子模型)，相当于在加边矩阵S中挑选相应行、列的元素构成加边的子矩阵。可能的子模型总个数是

$$C_n^1+C_n^2+\cdots+C_n^n=2^n-1$$

因此，如果要从n个自变元中挑选出对因变元影响最大的一些变元并建立回归模型，就需要用2^n-1个回归模型逐个拟合，然后依某种判别准则挑选出最优模型。显然当n较大时计算量相当大，甚至是难于实现的。

逐步回归方法〔5〕是解决上述问题较为得力的工具，但是按逐步回归使用F-检验方法对模型进行检验时，如果置信水平选取不当，可能导致误判。另外在经典回归分析中要求因变量z_1，z_2，…，z_N是不相关的，而在时序分析中，不同时刻因变元数据往往密切相关，所以不宜使用F-检验。因此在模型检验时通常采用AIC准则或BIC准则，以使准则函数达到极小的模型作为最佳模型。

由回归分析可知〔5〕，如果我们从某个已有的模型出发，要从现有的变量中剔除一个变量，或者是引入一个新的变量，并不需要对待选变元逐个求解线性方程组，而只需对加边矩阵进行一次消去变换，很容易便得到变化后模型的残差平方和及参数，这样就可以通过较少的计算得出较好的模型。对矩阵$(a_{ij})_{n\times n}$进行(k, k)消去变换定义如下（a'_{ij}表变换后的元素）

$$a'_{ij}=\begin{cases} 1/a_{kk} & i=k,\ j=k \\ a_{kj}/a_{kk} & i=k,\ j\neq k \\ -a_{ik}/a_{kk} & i\neq k,\ j=k \\ a_{ij}-a_{ik}a_{kj}/a_{kk} & i\neq k,\ j\neq k \end{cases} \qquad (8\text{-}2\text{-}20)$$

关于消去变换的性质可参阅〔5〕中第一章。

用近似AIC（或BIC）准则挑选回归变元的中心思想是：从高阶模型到低阶模型逐步剔除不重要变元，使准则函数值为最小的模型便是近似最佳模型

用T_i表示对矩阵施行(i, i)消去变换。我们先考察最高阶——n阶模型。为了得到n阶模型（8-2-17）式的参数估计$\hat{\alpha}$和模型残差，我们对初始加边矩阵（8-2-18）式连续进行n次消去

变换，为简化符号，变换后的矩阵及诸元素仍用原符号表示

$$S \Leftarrow T_n(T_{n-1}(\cdots(T_1 S)\cdots))$$

最后得到的变换后矩阵第$n+1$列的前n个元素$y_1, \cdots, y_n$就是n阶标准化模型的参数$\alpha_i^{(n)\prime}(i=1,2,\cdots,n)$。令$\widehat{\sigma_z}^2$，$\widehat{\sigma_i}^2(i=1,\cdots,n)$分别表示$\{z_t, 1\leqslant t\leqslant N\}$，$\{x_{it}, 1\leqslant t\leqslant N\}(i=1,\cdots,n)$的样本方差，$(\widehat{\sigma_\varepsilon}^{(n)})^2$表$n$阶模型的拟合残差方差，用$\alpha_i^{(n)}$，$i=1,\cdots,n$表$n$阶模型参数，则应有如下关系

$$\alpha_i^{(n)} = \alpha_i^{(n)\prime}\sqrt{\widehat{\sigma_z}^2/\widehat{\sigma_i}^2}$$

$$(\widehat{\sigma_\varepsilon}^{(n)})^2 = \widehat{\sigma_z}^2 y_0$$

y_0是变换后S矩阵的右下角元素。因此，从消去变换后的加边矩阵中，很容易将相应模型的参数和残差方差求出，相应的AIC和BIC值为

$$\text{AIC}_n = \log(\widehat{\sigma}_\varepsilon^{(n)})^2 + 2n/N$$

$$\text{BIC}_n = \log(\widehat{\sigma}_\varepsilon^{(n)})^2 + n\log N/N$$

由回归分析理论知，因变元z_t的方差可分解为

$$\sum_{t=1}^{N}(z_t-\bar{z})^2 = \sum_{t=1}^{N}(z_t-\widehat{z_t})^2 + \sum_{t=1}^{N}(\widehat{z_t}-\bar{z})^2 \quad (8\text{-}2\text{-}21)$$

其中$\bar{z}$是z_t的均值，$\widehat{z}$是z_t的某种估计值。式中右边第一项是残差平方和，它体现了除了自变量以外其它随机因素对z_t的影响；第二项是回归平方和，它体现了估计量$\widehat{z_t}$涉及到的自变元对z_t变差的影响，如果我们采用的自变量个数越多则对z_t变差的总贡献越大。如果从全部考虑的因素中去掉一个因素，则回归平方和只会减少，不会增加。减少的数值越少，说明该因素对回归平方和的贡献越少，即该因素越不重要。取消某一自变量后回归平方和减少的数值称为因变量z_t对这个变量的偏回归平方和。如果我们从

模型（8-2-7）的n个自变量中剔除第i个，则相应的偏回归平方和可由下式计算：

$$p_{n-1}^{(i)}=y_i^2/s_{ii} \qquad 1\leqslant i\leqslant n \qquad (8\text{-}2\text{-}22)$$

p的下标$n-1$表示所考虑模型中有$n-1$个自变量。

取
$$p_{n-1}^{(i_n)}=\min_{1\leqslant i\leqslant n} p_{n-1}^{(i)}$$

说明在n个自变量中，第i_n个变量（i_n是1与n之间的某个整数）偏回归平方和最小，即它对回归平方和的贡献最小，因此它对因变量的作用最小，可以将它剔除。这相当于在矩阵S中把第i_n行、第i_n列删除。

为了减少消去变换的计算量，我们对矩阵S的行做如下调整：

第i_n行

↙　　↖

第$n+1$行→第n行

调整后的矩阵的第i_n行是原矩阵第n行元素，原来的第$n+1$行移到了第n行的位置上，而原来的第i_n行移到第$n+1$行。对矩阵的列也做相应调整，调整的结果是把要剔除的那行、列换到最外层，即第$n+1$行，$n+1$列处。调整后矩阵仍记作S，然后对S施行（$n+1$，$n+1$）消去变换，进行变换的效果就是消去了欲剔除的变元。变换后的矩阵仍记作S，S的第n列中前$n-1$个元素y_1，…，y_{n-1}就是$n-1$阶标准化模型中残差平方和最小的那个标准化模型的参数，S的（n，n）元素y_n就是该模型的残差平方和。我们将n-1阶模型中残差最小者的AIC值（BIC值）算出，并与前面已得到的n阶模型的AIC值）BIC值）进行比较以决定取舍，然后对$n-1$个$n-2$阶模型的偏回归平方和进行比较，调整行、列，（注意这时是按：第n行（列）⇐第i_{n-1}行(列)，第$n-1$行(列)⇐第n行（列），第i_{n-1}行（列）⇐第$n-1$行（列）），进行消去

变换，删除该变元……，依此类推，每做一步模型减少一阶，同时进行消去变换的范围也减少一行、一列，一直做到零阶模型。我们对各阶模型的AIC(BIC) 值进行比较。假定p阶模型的AIC(BIC)值最小，便确定p阶模型为最佳，相应参数为$\hat{\alpha}_0, \hat{\alpha}_{i1}, \cdots, \hat{\alpha}_{i_p}$，模型为

$$z_t = \hat{\alpha}_0 + \hat{\alpha}_{i1} x_{i_1 t} + \cdots + \hat{\alpha}_{i_p} x_{i_p t} + \varepsilon_t \qquad (8\text{-}2\text{-}23)$$

上述方法是一个从高阶到低阶剔除变元的过程。

当然我们也可以低阶到高阶逐个选入变元，以建立最佳模型。当样本个数趋于无穷时，前一种方法挑选的最佳模型接近于真实模型，而后一种方法，挑选的最佳模型参数个数可能超过真实模型。有关证明请参阅〔22〕。

前面已经指出，要想拟合形如（8-2-17）式模型（阶数不超过n），并按照AIC（BIC）准则确定最优模型，需要拟合模型的个数是2^n-1，而依照前述从高阶到低阶逐个剔除变元的方法，需检验模型的总个数是$\frac{n(n-1)}{2}+1$，只是全部模型中的一部分，当然这部分模型综合了自变量中最重要的信息，但得到的最优模型是近似最优的，我们相应地称所使用的准则为近似最小信息量准则，虽然我们在模型最优化程度上做了些让步，但计算量却大大减少，从而使得在有些情形下几乎无法实现的问题，得以近似地实现。

用本节所讲的挑选回归变元的方法确定的模型往往是疏系数的，即在模型（8-2-23）中下标i_1，i_2，…，i_p通常是{1，2，3，…，}的不连贯的子集，下节的实例中提供了更具体的说明。挑选变元建立疏系数模型的具体实现可调用程序SMLR。

4． 实例分析 本节中我们将结合实例进一步阐明如何挑选

变元、建立线性混合回归模型。例1中通过对仿真数据的分析，说明从高阶到低阶剔除变元进行建模的步骤；在例2中对某地化肥逐年产量数据建立疏系数自回归模型，并进行预报；例3是对多元气象资料建立线性混合回归模型；例4中对地震地形变资料及温度、雨量资料建立增广混合回归模型，并对自变元进行挑选。

例1 仿真数据的分析。我们按照模型

$$x_t=0.6x_{t-2}+\varepsilon_t$$

调用程序DAGENT产生随机序列$\{x_t, 1\leqslant t\leqslant N\}$，N=200采用自回归模型

$$x_n=a_0+a_1x_{t-1}+\cdots+a_nx_{t-n}+\varepsilon_t$$

对该序列进行拟合，最大迟后取为$n=3$。调用建立疏系数线性混合回归模型的子程序SMLR，并在程序元件SUBLSM中取

$$a_{ti}\Leftarrow x_{t-i}, \quad z_t\Leftarrow x_t, \quad 1\leqslant i\leqslant 3, \quad 1\leqslant t\leqslant 200。$$

得到和正规方程（8-2-9）相应的加边矩阵

$$S\Leftarrow\begin{pmatrix}1.0000 & 0.2164 & 0.5602 & 0.2261\\ 0.2164 & 1.0000 & 0.2081 & 0.5532\\ 0.5602 & 0.2081 & 1.0000 & 0.1690\\ 0.2261 & 0.5532 & 0.1690 & 1.0000\end{pmatrix}$$

式（8-2-8）、（8-2-9）中进行标准化所用到的有关量是

$$\bar z=-0.0472, \quad \bar a_1=-0.0482, \quad \bar a_2=-0.0392, \bar a_3=-0.0508$$

$$\sqrt{s_{zz}}=16.5671, \quad \sqrt{s_{11}}=16.5850, \quad \sqrt{s_{22}}=16.5483,$$

$$\sqrt{s_{33}}=16.7424$$

为得到最高阶自回归模型（$n=3$），对前面矩阵S进行3次消去变换

$$S\Leftarrow T_3(T_2\ (T_1S)\ )$$

变换后得到的矩阵是

$$S \Leftarrow \begin{pmatrix} 1.4797 & -0.1543 & -0.7967 & 0.1145 \\ -0.1543 & 1.0614 & -0.1345 & 0.5295 \\ -0.7967 & -0.1345 & 1.4743 & -0.0054 \\ -0.1145 & -0.5295 & 0.0054 & 0.6821 \end{pmatrix}$$

从该矩阵S中的第四列元素可得到标准化模型参数

$$\alpha_1^{(3)\prime}=0.1145,\ \alpha_2^{(3)\prime}=0.5295,\ \alpha_3^{(3)\prime}=-0.0054$$

这些参数经变换后可得到$\alpha_0^{(3)}$，$\alpha_1^{(3)}$，$\alpha_2^{(3)}$，$\alpha_3^{(3)}$，因而3阶模型为

$$x_t=-0.0212+0.1144x_{t-1}+0.5301x_{t-2}-0.0053x_{t-3}+\varepsilon_t$$

由矩阵S右下角元素可计算出3阶模型的残差方差为

$$(\hat{\sigma}_\varepsilon^{(3)})^2=16.5671^2\times 0.6821/197=0.9503$$

BIC值为

$$\text{BIC}_3=\log(\hat{\sigma}_\varepsilon^{(3)})^2+2\log(N-n)\times 3/(N-n)$$

由式（8-2-22）可求得偏回归平方和是

$$p_2^{(1)}=0.0089,\ p_2^{(2)}=0.2642,\ p_2^{(3)}=0.00002.$$

其中$p_2^{(3)}$最小，故应将第3个变元x_{t-3}剔除。把前述S阵的第3、4行元素交换，然后把第3、4列元素交换，得到矩阵

$$\begin{pmatrix} 1.4797 & -0.1543 & 0.1145 & -0.7967 \\ -0.1543 & 1.0614 & 0.5295 & -0.1345 \\ -0.1145 & -0.5295 & 0.6821 & 0.0054 \\ -0.7967 & -0.1345 & -0.0054 & 1.4743 \end{pmatrix}$$

对该矩阵施行（4，4）消去变换，得到

$$S \Leftarrow \begin{pmatrix} 1.0491 & -0.2270 & 0.1116 & 0.5404 \\ -0.2270 & 1.0491 & 0.5290 & 0.0912 \\ -0.1116 & -0.5290 & 0.6821 & -0.0037 \\ -0.5404 & -0.0912 & -0.0037 & 0.6783 \end{pmatrix}$$

二变元模型的参数可由该阵第3列前两元素换算得到。二变元模型是

$$x_t = -0.0211 + 0.1115 x_{t-1} + 0.5296 x_{t-2} + \varepsilon_t$$

与前述类似可以计算出

$$(\hat{\sigma}_\varepsilon^{(2)})^2 = 0.9503 \qquad \mathrm{BIC}_2 = 0.05637$$

同样可算出

$$p_1^{(1)} = 0.0119, \qquad p_1^{(2)} = 0.2668$$

其中$p_1^{(1)}$较小，故应剔除变元x_{t-1}，在前述S中进行行、列调换：

$$\begin{array}{cc} \text{第1行} & \text{第1列} \\ \swarrow \quad \nwarrow & \swarrow \quad \nwarrow \\ \text{第3行}\rightarrow\text{第2行} & \text{第3列}\rightarrow\text{第2列} \end{array}$$

得到

$$\begin{pmatrix} 1.0491 & 0.5290 & -0.2270 & 0.5404 \\ -0.5290 & 0.6821 & -0.1116 & 0.0912 \\ -0.2270 & 0.1116 & 1.0491 & -0.0037 \\ -0.5404 & -0.0912 & -0.0037 & 0.6783 \end{pmatrix}$$

对该矩阵左上角的3维子矩阵施行（3,3）消去变换得到

$$\begin{pmatrix} 1.0000 & 0.5532 & 0.2164 & 0.5404 \\ -0.5532 & 0.6940 & 0.1064 & 0.0912 \\ -0.2164 & 0.1064 & 0.9532 & -0.0037 \\ -0.5404 & -0.0912 & -0.0037 & 0.6787 \end{pmatrix}$$

一变元模型为

$$x_t = -0.0255 + 0.5538 x_{t-2} + \varepsilon_t$$

该模型的残差方差及BIC值为

$$(\hat{\sigma}_\varepsilon^{(1)})^2 = 0.9669 \qquad \mathrm{BIC}_1 = 0.0199$$

比较各阶模型的BIC值可知，BIC_1最小，故一变元模型为最佳，且其参数估计值与产生随机序列的真实模型十分接近。

从以上叙述我们可以详细了解建立近似最佳疏系数混合线性回归模型的具体步骤。

例2. 化肥产量数据的建模与预测　某地化肥逐年产量的数

据列在表8-7中（1949年—1981年），我们要用为数不多的观察值（$N=33$）建立模型，并外推预报1982年，1983年…，2000年各年的产量。我们用$\{x_t, 1\leqslant t\leqslant N\}$表示观察序列。

由表8-7及图8-11中看到，化肥产量数据具有明显的增长趋势，而且数据量较少，预报的步数又相当长，因此用多项式回归或直接用自回归模型都是不恰当的，前者的预报方差往往很大，而后者只适用于平稳观察序列。按照8.1节所述方法，对具有增长趋势的序列进行差分处理使之平稳化，即令

$$y_t = x_{t+1} - x_t \qquad 1\leqslant t\leqslant 32$$

若对$\{y_t\}$建立自回归模型，可能导致需用较高阶数的模型进行拟合。在观察数据较少，而待估计参数又较多的情况下，势必影响模型及预报的精度。因此采用疏系数自回归模型进行拟合，并用近似BIC准则挑选变元，这样可以减少计算量并节省参数，从而保证参数估计的相对稳定性和提高预报的可靠性。

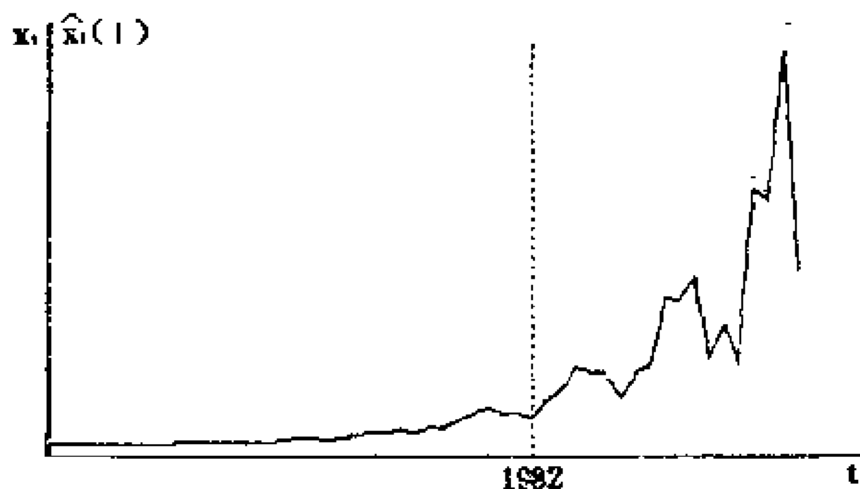


图8－11　化肥产量及预报值

我们考虑形如

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_n y_{t-n} + \varepsilon_t$$

的模型，最大迟后取$n=8$。调用程序SMLR，并令

表8—7 化肥产量及预报值(单位,万吨)

1949	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959
1.3	3.9	6.9	8.3	11.5	18.7	23.2	31.6	35.0	53.5	46.5
1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970
44.0	40.5	59.9	98.6	99.6	155.3	206.3	198.1	171.9	218.8	286.7
1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981
324.6	384.9	424.8	381.5	469.5	475.2	675.3	900.9	1037.6	938.0	924.8
1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992
790.2	1256.4	1580.0	2148.5	2051.9	2032.6	1350.5	2017.8	2270.5	4150.4	4069.4
1993	1994	1995	1996	1997	1998	1999	2000			
4716.0	2464.1	3376.9	2306.7	7226.9	6903.5	11194.5	4920.2			

$$a_{ti} \Leftarrow y_{t-i}, \quad z_t \Leftarrow y_t, \quad 1 \leqslant i \leqslant 8, \quad 1 \leqslant t \leqslant 32$$

得到近似最佳模型是

$$y_t = 9.975 + 1.316 y_{t-6} - 0.651 y_{t-3} + 1.454 y_{t-8} + \varepsilon_t$$

将$y_t = x_{t+1} - x_t$代入上式得到x_t的近似最佳模型

$$x_{t+1} = 9.975 + x_t + 1.316(x_{t-7} - x_{t-6}) - 0.651(x_{t-4} - x_{t-3}) + 1.454\ (x_{t-9} - x_{t-8}) + \varepsilon_t$$

x_t的预报值可由下式递推算出

$$\hat{x}_{t+1} = 9.975 + \hat{x}_t + 1.316(\hat{x}_{t-7} - \hat{x}_{t-6}) - 0.651(\hat{x}_{t-4} - \hat{x}_{t-3}) + 1.454(\hat{x}_{t-9} - \hat{x}_{t-8}) \qquad t \geqslant N$$

我们约定在该式中当$t \leqslant N$时，$\hat{x}_t = x_t$，而当$t > N$时$\hat{x}_t$用预报值代替。表8-7及图8-11中还给出了1982—2000年化肥产量的预测值，为该地区经济远景规划中化肥的预期产量提供了科学依据。

例3 多元气象资料的统计分析 这里用多变量混合回归模型对7.4.4节中提到的气象资料（数据见表7-2）进行拟合。对每一变元$\{x_{it}\}$ ($i=1,2,3,4$) 建立模型式 (8.2.12)，并用近似AIC准则挑选变元，即不仅确定出模型中包含哪些气象要素，而且要选出某种变元的哪些迟后步数起重要作用。换句话说，也即式（8-2-12）中应包含哪些$x_{i,\ t-l}$。由于气象资料的月平均数据存在12个月的周期，我们在式(8-2-12)中取最大迟后$p=15$，变量个数$s=4$。如果采用满系数模型，待估参数总数是$4\times(4\times15)=240$，这样多的参数要从4×120个数据中估计，其相对稳定性必然很差。因此应当按照式（8-2-12），对每一分量用疏系数线性混合回归进行拟合。此外，若使用一般的AIC准则，所需计算相当于$(2^{15}-1)^4$个线性回归模型拟合的计算量，而使用近似的AIC准则及消去变换，可将计算量降低到相当于拟合4×60个线性回归模型。具体计算结果表明，这样建立的多变量疏系数线性混合回归模型可以保证参数估计值的统计稳定性。最后得到的拟合结果如下：

$$
\begin{aligned}
x_{1t} &= 62.6 - 9.15x_{2,t-8} + 8.52x_{3,t-1} + 2.18x_{4,t-11} \\
&\quad + \varepsilon_t^{(1)} \\
x_{2t} &= 23.8 + 0.264x_{2,t-2} - 0.281x_{2,t-6} - 0.301x_{2,t-9} \\
&\quad + 0.464x_{3,t-3} - 0.879x_{3,t-9} - 0.457x_{3,t-13} \\
&\quad - 0.267x_{3,t-14} + 0.582x_{4,t-9} + 0.355x_{4,t-14} + \varepsilon_t^{(2)} \\
x_{3t} &= 12.1 + 0.00499x_{1,t-3} - 0.00545x_{1,t-15} \\
&\quad - 0.387x_{2,t-3} + 0.331x_{2,t-8} + 0.398x_{2,t-9} + \varepsilon_t^{(3)} \\
x_{4t} &= -0.831 + 0.00549x_{1,t-8} + 0.00683x_{1,t-10} \\
&\quad - 0.00837x_{1,t-15} + 0.322x_{2,t-6} + 0.344x_{2,t-10} \\
&\quad - 0.319x_{3,t-7} + 0.458x_{3,t-9} + 0.346x_{3,t-10} \\
&\quad - 0.451x_{4,t-9} + 0.359x_{4,t-12} + \varepsilon_t^{(4)}
\end{aligned}
$$

以上共31个参数，式中$15<t\leqslant 120$。我们用矢量表示式(8-2-13)

$$
X_t = a_0 + \sum_{i=1}^{15} A_i X_{t-i}
$$

对照前面四个模型可知

$$
a_0 = (62.6, 23.8, 12.1, -0.831)^T
$$

$$
A_1 = \begin{pmatrix} 0 & 0 & 8.52 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad
A_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0.264 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},
$$

$$
A_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0.464 & 0 \\ 0.00499 & -0.387 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad
A_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},
$$

$$A_5=\begin{pmatrix}0&0&0&0\\0&0&0&0\\0&0.331&0&0\\0&0&0&0\end{pmatrix},\quad A_6=\begin{pmatrix}0&0&0&0\\0&-0.281&0&0\\0&0&0&0\\0&0.322&0&0\end{pmatrix},$$

$$A_7=\begin{pmatrix}0&0&0&0\\0&0&0&0\\0&0&-0.353&0\\0&0&-0.319&0\end{pmatrix},\quad A_8=\begin{pmatrix}0&9.15&0&0\\0&0&0&0\\0&0&0&0\\0.00549&0&0&0\end{pmatrix},$$

$$A_9=\begin{pmatrix}0&0&0&0.582\\-0.301&0&-0.879&0\\0&0.398&0.342&0\\0&0&0.458&-0.451\end{pmatrix},$$

$$A_{10}=\begin{pmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0.00683&0.344&0.346&0\end{pmatrix},$$

$$A_{11}=\begin{pmatrix}0&0&0&2.18\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix},\quad A_{12}=\begin{pmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0.359\end{pmatrix},$$

$$A_{13}=\begin{pmatrix}0&0&0&0\\0&0&-0.457&0\\0&0.00499&0.303&0\\0&0&0&0\end{pmatrix},\quad A_{14}=\begin{pmatrix}0&0&0&0\\0&0&-0.267&0.355\\0&0&0&0\\0&0&0&0\end{pmatrix},$$

$$A_{15}=\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -0.00545 & 0 & 0 & 0 \\ -0.00837 & 0 & 0 & 0 \end{pmatrix}.$$

用 $\hat{\sigma}_i^0$ 表示观察数据 $\{x_{it},\ t=1,2,\cdots\}$ 的样本均方差，用 $\hat{\sigma}_i$ 表示前述模型的拟合均方差，本例的拟合结果见表8—8。

表8—8 多元气象资料拟合结果

i	1	2	3	4
$\hat{\sigma}_i^0$	65.516	7.958	8.637	8.796
$\hat{\sigma}_i$	53.015	1.435	1.257	1.330
$\hat{\sigma}_i/\hat{\sigma}_i^0$	0.8092	0.1803	0.1455	0.1512

从表中可以看到，$i=2,3,4$ 时 $\hat{\sigma}_i/\hat{\sigma}_i^0$ 的值不超过0.19，而 $\hat{\sigma}_1/\hat{\sigma}_2^0=0.8092$，这说明水量的拟合效果是最差的，通过建立线性混合回归模型，只将水量的均方差压缩了20%左右。因此，要改进对水量的拟合，需要采用其它方法，如增加数据采样频率，对数据进行预处理，采用门限回归模型等等，这里不再做进一步探讨。

例4 地形变与气温、雨量增广混合回归模型的建立。某地震台站积累了大量地震观察资料：从1970年1月1日到1980年8月31日地形变、雨量、气温的日观测值，它们分别用 z_t，x_{1t}，x_{2t} 表示，每种数据的个数都是 $N=4261$。图8—12中给出了每种观察数据随时间变化的曲线。

这里需要建立反映地形变变化规律的统计模型。由图8-12看到，地形变数据有明显的年周期变化，起伏趋势基本上与气温、雨量变化相符。因此考虑采用增广线性混合回归模型进行拟合，即通过地形变对气温雨量的回归去掉地形变数据的季节趋势（年

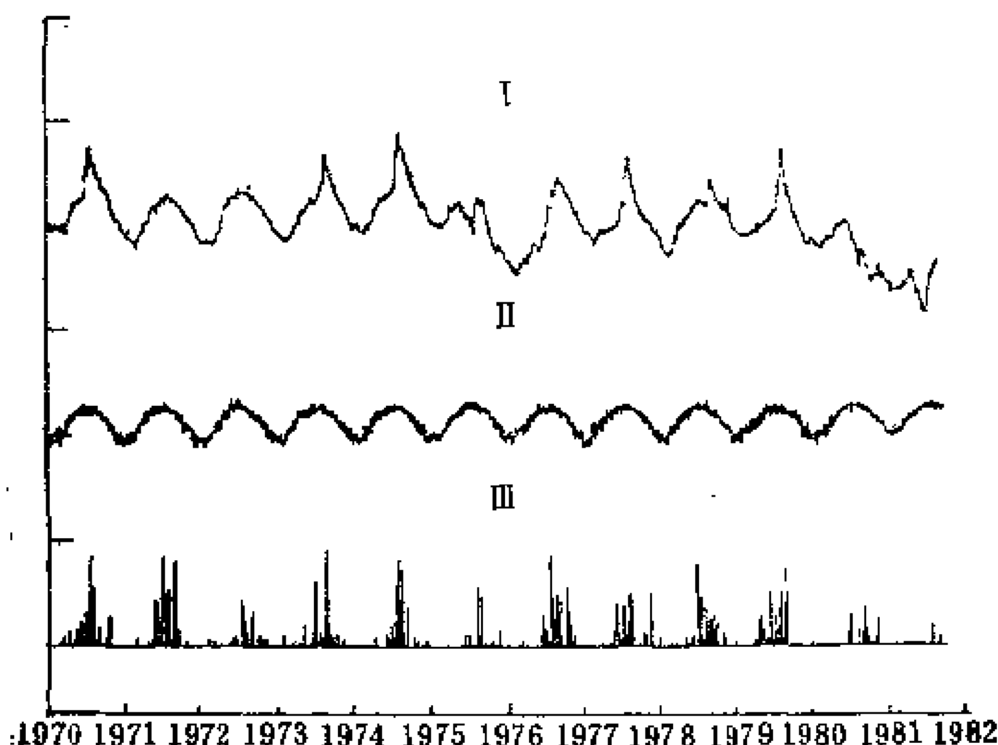


图8—12 地震量测数据曲线

（Ⅰ，地形变 Ⅱ，气温 Ⅲ，雨量）

周期），然后对残量作自回归拟合。

$$z_t=\beta_0+\beta_1^{(1)}x_{1,t-1}+\cdots+\beta_{q_1}^{(1)}x_{1,t-q_1}+\beta_1^{(2)}x_{2,t-1}$$
$$+\cdots+\beta_{q_2}^{(2)}x_{2,t-q_2}+\varepsilon_t$$

$$\varepsilon_t=\alpha_0+\alpha_1\varepsilon_{t-1}+\cdots+\alpha_{\bar{p}}\varepsilon_{t-\bar{p}}+\eta_t$$

由于数据量相当大，而且1976年7月28日唐山地震前后地形变资料发生明显变异，因此我们应对原始数据进行适当平均处理，并且分段地用增广线性回归模型拟合。对原观察序列每5点平均得到一点，于是得到处理后的序列 $\{z'_t,\ x_1{}'_t,\ x_2{}'_t,\ t=1,2,\cdots,N'\}$，$N'=N/5$。对处理后的序列采用增广混合回归模型拟合，样本长度取 $N'_1=300$。用8.2.2节中讲过的线性最小二乘方法拟合增广混合回归模型，对前300个数据的拟合结果是

$$z'_t=494.73+0.45009x'_{1,t-4}+0.52338x'_{1,t-6}$$

$$+0.4613x'_{1,t-12}+0.36931x'_{2,t-4}+3.00109x'_{2,t-8}$$
$$+0.29885x'_{2,t-11}+0.39351x'_{2,t-12}+\varepsilon_t$$
$$\varepsilon_t=-0.0087205+0.60489\varepsilon_{t-1}+0.30133\varepsilon_{t-2}+\eta_t$$

该模型的拟合残差的均方差分别是 $\sigma_\varepsilon=7.7645$，$\sigma_\eta=3.6918$，而该段数据 $\{z^1_t\}$ 的均方差 $\sigma_{z1}=54.945$，可见拟合残差均方缩小到不及原来的十四分之一，拟合效果显著。

将序列 $\{z^1_t,\ x^1_{1t},\ x^1_{2t},\ 1\leqslant t\leqslant 800\}$ 每次向右移动5点，重复进行增广混合线性回归模型拟合，并对各个模型进行分析，可以看出地震前后模型统计特征有明显的变化（表现在模型参数及残差的改变），由此可划分地震前的正常期、地震时的异常期和地震后的正常期。所以上述方法对于分析地形变观察资料并从中提取地震异常信息是可取的.

8.3 门限回归（自回归）模型

门限自回归模型（*Threshold Autoregressive Model*）由汤家豪于1978年首先提出[23]，近几年来对这类模型的研究日益深入完善，在应用上也很有成效，并进一步扩展到回归模型中。

门限回归模型可用来解决一类非线性问题，其基本思想是：把非线性模型按照某一变元的不同取值范围，采用若干个线性模型来描述，另外还将微分方程中极限环的概念引入非线性随机系统，有效地描述具有周期规律的过程，而且由于门限的控制作用，保证了模型的稳定性。此外，这类模型还可以做为突变现象的一种描述手段。

本节将介绍门限回归、门限自回归、门限混合回归等模型，并给出实现步骤及一些应用的实例。

1.门限回归（自回归）模型的定义 在引入定义之前，我们分析一个简单的例子。我们考虑某种商品（如电视机）的逐年销售量z_t，假定它与该商品逐年售价x_{1t}，居民逐年平均收入x_{2t}，居民逐年基本生活消费x_{3t}等因素有关。我们可以用一个线性回归模型来描述

$$z_t=\alpha_0+\alpha_1x_{1t}+\alpha_2x_{2t}+\alpha_3x_{3t}+\varepsilon_t$$

然而事实上，当居民的年节余低于或超过该商品售价时，商品的销售情况是截然不同的，如果年节余超过该商品售价时，居民几乎随时可以购买。因此用一个统一的线性回归模型拟合，效果是不会好的，它没有能反映出x_{2t}低于或超过某一值时，对z_t影响的突变规律。很自然地，应当考虑按x_{2t}的不同取值，采用不同的线性回归模型。具体地说，若d年前居民平均收入低于或超过某一标准，则该商品的销售量将呈现不同的变化规律，用公式表示就是

$$z_t=\begin{cases}\alpha_0^{(1)}+\alpha_1^{(1)}x_{1t}+\alpha_2^{(1)}x_{2t}+\alpha_3^{(1)}x_{3t}+\varepsilon_t^{(1)} & \text{当}x_{2,t-d}\leqslant\bar{x}\\ \alpha_0^{(2)}+\alpha_1^{(2)}x_{1t}+\alpha_2^{(2)}x_{2t}+\alpha_3^{(2)}x_{3t}+\varepsilon_t^{(2)} & \text{当}x_{2,t-d}>\bar{x}\end{cases}$$

其中d是迟后量，$\bar{x}$是门限值。

设$\{z_t\}$是因变元观察序列，$\{x_{it}\ 1\leqslant i\leqslant s\}$是自变量观察序列。我们引入下列定义

定义1 门限回归模型定义为

$$z_t=\begin{cases}\alpha_0^{(1)}+\alpha_1^{(1)}x_{1t}+\cdots+\alpha_s^{(1)}x_{st}+\varepsilon_t^{(1)} & \text{当}x_{i,t-d}\leqslant\bar{x}_1\\ \alpha_0^{(2)}+\alpha_1^{(2)}x_{1t}+\cdots+\alpha_s^{(2)}x_{st}+\varepsilon_t^{(2)} & \text{当}\bar{x}_1<x_{i,t-d}\leqslant\bar{x}_2\\ \vdots \\ \alpha_0^{(r)}+\alpha_1^{(r)}x_{1t}+\cdots+\alpha_s^{(r)}x_{st}+\varepsilon_t^{(r)} & \text{当}\bar{x}_{r-1}<x_{i,t-d}<+\infty\end{cases}\quad(8\text{-}3\text{-}1)$$

其中自变元序列$\{x_{it}\}$称为门限变元，$\bar{x}_1$，$\bar{x}_2$，$\cdots\bar{x}_{r-1}$为门限值，d为门限元迟后量。

定义2 门限自回归模型。设$\{z_t\}$是单变量观察序列，门限自回归模型定义为

$$z_t=\begin{cases}\beta_0^{(1)}+\beta_1^{(1)}z_{t-1}+\cdots+\beta_{p1}^{(1)}z_{t-p_1}+\varepsilon_t^{(1)} & \text{当}z_{t-d}\leqslant\bar{z}_1\\ \beta_0^{(2)}+\beta_1^{(2)}z_{t-1}+\cdots+\beta_{p2}^{(2)}z_{t-p_2}+\varepsilon_t^{(2)} & \text{当}\bar{z}_1<z_{t-d}\leqslant\bar{z}_2\\ \vdots \\ \beta_0^{(r)}+\beta_1^{(r)}z_{t-1}+\cdots+\beta_{pr}^{(r)}z_{t-p_r}+\varepsilon_t^{(r)} & \text{当}\bar{z}_{r-1}<z_{t-d}\end{cases}\quad(8\text{-}3\text{-}2)$$

其中$\bar{z}_1$，$\bar{z}_2$，$\cdots$，$\bar{z}_{r-1}$为门限值，d是门限迟后量。

定义3 门混混合回归模型。设$\{z_t\}$及$\{x_{it},1\leqslant i\leqslant s\}$分

别为因变量及自变量观察序列，门限混合回归模型定义为

$$z_t=\begin{cases}\beta_0^{(1)}+\beta_1^{(1)}z_{t-1}+\cdots+\beta_{q1}^{(1)}z_{t-q1}\\ +\sum\limits_{j=1}^{s}\sum\limits_{i=0}^{p_j^{(1)}}\alpha_{ji}^{(1)}x_{j,t-i} \quad 当x_{l,t-d}\leqslant\bar{x}_1\\ \beta_0^{(2)}+\beta_1^{(2)}z_{t-1}+\cdots+\beta_{q2}^{(2)}z_{t-q2}\\ +\sum\limits_{j=1}^{s}\sum\limits_{i=0}^{p_j^{(2)}}\alpha_{ji}^{(2)}x_{j,t-i} \quad 当\bar{x}_1<x_{l,t-d}\leqslant\bar{x}_2\\ \vdots\\ \beta_0^{(r)}+\beta_1^{(r)}z_{t-1}+\cdots+\beta_{qr}^{(r)}z_{t-qr}\\ +\sum\limits_{j=1}^{s}\sum\limits_{i=0}^{p_j^{(r)}}\alpha_{ji}^{(r)}x_{j,t-i} \quad 当\bar{x}_{r-1}<x_{l,t-d}\end{cases} \quad (8\text{-}3\text{-}3)$$

其中$\bar{x}_1$，$\bar{x}_2$，…，$\bar{x}_{r-1}$为门限值，$\{x_{l,t},1\leqslant t\leqslant N\}$为门限变元，$d$是门限迟后。当然也可以取因变元$z_t$作门限元，门限值为$z_1$，$z_2,\cdots,z_{r-1}$，式（8-3-3）的条件应相应改变。

由以上定义可以看出式（8-3-1）和(8-3-2)都是式(8-3-3)的特例。式（8-3-2）中只包含门限自回归部分，而式（8-3-1）中只包含自变元的门限回归部分。

2.门限自回归模型的特性 门限自回归模型是具有较广泛意义的一种非线性模型。例如考察一阶非线性自回归模型

$$x_t=f(x_t-1)+\varepsilon_t$$

其中f是连续函数，且在区间〔a,b〕上为一致连续，于是对任何$x\in$〔a,b〕，利用近似展开，存在下面关系式

$$f(x)\cong f(\bar{x}_{j-1})+\frac{f(\bar{x}_j)-f(\bar{x}_{j-1})}{\bar{x}_j-\bar{x}_{j-1}}(x-\bar{x}_{j-1})=\alpha_{j-1}+\alpha_j x$$

其中$\alpha_j\equiv[f(\bar{x}_j)-f(\bar{x}_{j-1})]/[\bar{x}_j-\bar{x}_{j-1}]$，$\alpha_{j-1}\equiv f(\bar{x}_{j-1})-$

$a_j \tilde{x}_{j-1}$，$a=\tilde{x}_0<\tilde{x}_1<\cdots\tilde{x}_r=b$ 是〔a,b〕上的一组分割，分割的粗细由逼近精度决定。在给定精度下，一阶非线性回归模型可以由下式近似

$$x_t=a_{j-1}+a_j x_{t-1}+\varepsilon_t \qquad \tilde{x}_{j-1}\leqslant x_{t-1}<\tilde{x}_j$$

这正是一阶门限自回归模型，可见一阶非线性回归模型可由一阶门限自回归模型近似。

类似地，对于一般的k阶非线性自回归模型

$$x_t=f(x_{t-1},\cdots,x_{t-k})+\varepsilon_t$$

也可用高阶的门限自回归模型逼近．这说明了门限自回归模型的广泛意义。

为了进一步介绍门限自回归模型的特性，要简单了解微分方程极限环的概念。在第二章中已讨论过线性随机差分方程的结构和性质与对应的齐次差分方程有密切关系，如果齐次差分方程的解当$t\to\infty$时趋于零，则相应自回归模型是平稳的。那末非线性的门限自回归模型与相应的齐次差分方程有什么关系？后者的解具有什么性质？为叙述方便，令

$$p=\max(p_1,p_2,\cdots,p_r),$$

$$f(x_{t-1},x_{t-2},\cdots,x_{t-p})=$$

$$\begin{cases}\beta_0^{(1)}+\beta_1^{(1)}x_{t-1}+\cdots+\beta_{p_1}^{(1)}x_{t-p_1} & \text{当}x_{t-d}\leqslant\tilde{x}_1\\ \beta_0^{(2)}+\beta_1^{(2)}x_{t-1}+\cdots+\beta_{p_2}^{(2)}x_{t-p_2} & \text{当}\tilde{x}_1<x_{t-d}\leqslant\tilde{x}_2\\ \vdots & \\ \beta_0^{(r)}+\beta_1^{(r)}x_{t-1}+\cdots+\beta_{p_r}^{(r)}x_{t-p_r} & \text{当}\tilde{x}_{r-1}<x_{t-d}<+\infty\end{cases}$$

则与门限自回归模型相应的齐次差分方程为

$$x_t=f(x_{t-1},\ x_{t-2},\cdots,x_{t-p}) \qquad (8\text{-}3\text{-}4)$$

式（8-3-4）的解具有如下特性：

（1）若从某一组初始值$x_p,x_{p-1},\cdots,x_1$出发，式（8-3-4）给出的序列x_t满足$\lim\limits_{t\to\infty}x_t=x$，则称$x$为式（8-3-4）的极限点。

(2) 若从某组初始值$x_p, x_{p-1}, \cdots, x_1$出发，$x_t$渐近地与某一周期为$T$的周期序列相吻合，即

$$\lim_{t\to\infty}\sum_{i=1}^{T}(x_{t-i}-C_{t-i})^2=0$$

则称$\{C_t\}$是式(8-3-4)以T为周期的极限环。T是使C_t满足下式的最小周期

$$C_{t+T}=C_t \qquad t=1,2,3,\cdots$$

若从任何$(x_p, x_{p-1}, \cdots, x_1)$的邻域出发，$\{x_t\}$都与$\{C_t\}$渐近吻合，则称它为稳定极限环。

例如，考虑一阶齐次差分方程

$$x_t=\begin{cases}4x_{t-1} & 当\ |x_{t-1}|\leqslant\frac{1}{2}\\ \frac{1}{2}x_{t-1} & 当\ |x_{t-1}|>\frac{1}{2}\end{cases}$$

取初值$x_1=4$，则$x_2=2$，$x_3=1$，$x_4=\frac{1}{2}$，$x_5=2$，$x_6=1$，$x_7=\frac{1}{2},\cdots$，即$\{x_t\}$有极限环$\{\frac{1}{2},2,1,\frac{1}{2},2,1,\frac{1}{2},2,1,\cdots\}$。若取初值$x_1=5$，则$x_2=\frac{5}{2}$，$x_3=\frac{5}{4}$，$x_4=\frac{5}{8}$，$x_5=\frac{5}{16}$，$x_6=\frac{5}{4}$，$x_7=\frac{5}{8}$，$x_8=\frac{5}{16}$，$\cdots$，$\{x_t\}$具有极限环$\{\frac{5}{8}, \frac{5}{16}, \frac{5}{4}, \frac{5}{8}, \frac{5}{16}, \frac{5}{4}, \cdots\}$。然而对于一般的差分方程很难详尽举出它所有的极限点和极限环。

把微分方程极限环的概念引入随机的门限自回归模型中，是为了进一步探讨门限自回归模型的特性，及适合这类模型的周期序列与极限环的关系。有关这类问题的研究，目前仍在探索之中。下面举一个简单的例子，给出门限自回归模型与极限环的关系，说明在一定条件下，门限自回归模型可以用来描述含有较复杂周期规律的随机系统。

考虑一阶门限自回归模型

$$x_t=\begin{cases} a^{(1)}x_{t-1}+\varepsilon_t^{(1)} & x_{t-d}\leqslant\bar{x}_1 \\ a^{(2)}x_{t-1}+\varepsilon_t^{(2)} & \bar{x}_1<x_{t-d}\leqslant\bar{x}_2 \\ a^{(3)}x_{t-1}+\varepsilon_t^{(3)} & \bar{x}_2<x_{t-d}<+\infty \end{cases}$$

d是门限迟后量，模型参数$|a^{(1)}|<1$，$|a^{(3)}|<1$，$|a^{(2)}|>1$，$\bar{x}_1,\bar{x}_2$是门限值，且$\bar{x}_1<0<\bar{x}_2$。当递推值$x\in(-\infty,\bar{x}_1]$或$x\in(\bar{x}_2,+\infty]$时，递推是稳定的，相应差分方程有稳定解$x=0$，但是一旦递推值$x\in(\bar{x}_1,\bar{x}_2]$时，由于$|a^{(2)}|>1$，差分方程的解不稳定，因此递推发散；而当发散超出区间$(\bar{x}_1,\bar{x}_2]$进入$(-\infty,\bar{x}_1]$或$(\bar{x}_2,+\infty]$后，递推值又趋向于零。由上述可知，该门限自回归模型构成了周期性的稳定循环序列，因而具有极限环。

3.门限回归（自回归）模型的建立 由于门限回归(自回归)模型是按照门限元不同的取值范围而采取不同的拟合模型。为了建立模型方便，我们首先要把门限变元$\{x_t,\ t=1,2,\cdots,N\}$按照从小到大的顺序进行排列。当然，对于门限自回归模型需要排序的就是观察序列自身。与此同时，也要把因变量及其它量测分量也按相应关系进行调整。事实上，在计算机上实现时，为节省内存及调用方便，我们只需引入一个标号序列$\{i_t,1\leqslant t\leqslant N\}$，该序列描述了将门限元以小到大排列时，在原门限元序列中的取样顺序。为了说明标号序列的含义，我们举一个简单的例子。考察序列$\{x_t,1\leqslant t\leqslant 6\}$：

$$x_1=1,\ x_2=0,\ x_3=-1,\ x_4=3,\ x_5=-2,\ x_6=-3$$

门限值取为$\bar{x}_1=0.5$，在原序列中若按下面顺序取数

$$i_1=6,\ i_2=5,\ i_3=3,\ i_4=2,\ i_5=1,\ i_6=4$$

则取出的数就是按从小到大顺序排列的。因此如果有了标号序列$\{i_t,1\leqslant t\leqslant 6\}$，那么序列$\{x_{i_t},\ 1\leqslant t\leqslant 6\}$就是按从小到大顺序排列的。至于门限值$\bar{x}$则用$\{i_t,\ 1\leqslant t\leqslant 6\}$中相应的下标表示就是$t_1=4$。

设 $\{x_{l,t}, 1\leqslant t\leqslant N\}$ 是门限变元。通过比较序列中元素大小，得到相应的标号序列 $\{i_t, 1\leqslant t\leqslant N\}$，因此序列 $\{x_{l,i_t}, 1\leqslant t\leqslant N\}$ 是按从小到大顺序排列的，而 $\{x_{j,i_t}, 1\leqslant t\leqslant N, j\neq l\}$ 及 $\{z_{i_t}, 1\leqslant t\leqslant N\}$ 都是按照门限元从小到大排列顺序来取样的。

设 $t_0=-\infty$，$t_r=+\infty$，$t_1, t_2, \cdots, t_r$ 是门限值 $\bar{x}_1$，$\bar{x}_2$，$\cdots$，$\bar{x}_{r-1}$ 在序列 $\{x_{l,i_t}, 1\leqslant t\leqslant N\}$ 中标号 i_t 的下标 (t)，注意将下标推移 d 步，于是将门限回归模型（8-3-1）改写成

$$z_{i_t+d}=\begin{cases}\alpha_0^{(1)}+\alpha_1^{(1)}x_{1,i_t+d}+\cdots+\alpha_s^{(1)}x_{s,i_t+d}+\varepsilon_t^{(1)} \\ \qquad\qquad t_0<t\leqslant t_1 \\ \alpha_0^{(2)}+\alpha_1^{(2)}x_{1,i_t+d}+\cdots+\alpha_s^{(2)}x_{s,i_t+d}+\varepsilon_t^{(2)} \\ \qquad\qquad t_1<t\leqslant t_2 \\ \vdots \\ \alpha_0^{(r)}+\alpha_1^{(r)}x_{1,i_t+d}+\cdots+\alpha_s^{(r)}x_{s,i_t+d}+\varepsilon_t^{(r)} \\ \qquad\qquad t_{r-1}<t\leqslant t_r\end{cases} \tag{8-3-5}$$

门限自回归模型改写成

$$z_{i_t+d}=\begin{cases}\beta_0^{(1)}+\beta_1^{(1)}z_{i_t+d-1}+\cdots+\beta_{p_1}^{(1)}z_{i_t+d-p_1} \\ +\varepsilon_t^{(1)} \qquad\qquad t_0<t\leqslant t_1 \\ \beta_0^{(2)}+\beta_1^{(2)}z_{i_t+d-1}+\cdots+\beta_{p_2}^{(2)}z_{i_t+d-p_2} \\ +\varepsilon_t^{(2)} \qquad\qquad t_1<t\leqslant t_2 \\ \vdots \\ \beta_0^{(r)}+\beta_1^{(r)}z_{i_t+d-1}+\cdots+\beta_{p_r}^{(r)}z_{i_t+d-p_r} \\ +\varepsilon_t^{(r)} \qquad\qquad t_{r-1}<t\leqslant t_r\end{cases} \tag{8-3-6}$$

同理，门限混合回归模型可改写成

$$
z_{i_t+d}=\begin{cases}
\beta_0^{(1)}+\beta_1^{(1)}z_{i_t+d-1}+\cdots+\beta_{q_1}^{(1)}z_{i_t+d-q_1} \\
+\sum_{j=1}^{s}\sum_{i=0}^{p_j^{(1)}}a_{ji}^{(1)}x_{j,i_t+d-i} \quad \text{当} t_0<t\leqslant t_1 \\
\beta_0^{(2)}+\beta_1^{(2)}z_{i_t+d-1}+\cdots+\beta_{q_2}^{(2)}z_{i_t+d-q_2} \\
+\sum_{j=1}^{s}\sum_{i=0}^{p_j^{(2)}}a_{ji}^{(2)}x_{j,i_t+d-i} \quad \text{当} t_1<t\leqslant t_2 \\
\vdots \\
\beta_0^{(r)}+\beta_1^{(r)}z_{i_t+d-1}+\cdots+\beta_{q_r}^{(r)}z_{i_t+d-q_r} \\
+\sum_{j=1}^{s}\sum_{i=0}^{p_j^{(r)}}a_{ji}^{(r)}x_{j,i_t+d-i} \quad \text{当} t_{r-1}<t\leqslant t_r
\end{cases}
$$

如果门限分点的位置、个数及门限元迟后量给定，那么问题便归结为如何建立与一对门限值t_a，t_b（$t_1\leqslant t_a<t_b\leqslant t_{r-1}$）相对应的模型

在（t_a，t_b〕上模型（8-3-5）可写成

$$
z_{i_t+d}=a_0+a_1x_{1,i_t+d-1}+\cdots+a_sx_{s,i_t+d}+\varepsilon_t \qquad t_a<t\leqslant t_b \tag{8-3-8}
$$

（8-3-6）在（t_a,t_b〕的模型是

$$
z_{i_t+d}=\beta_0+\beta_1z_{i_t+d-1}+\cdots+\beta_pz_{i_t+d-p}+\varepsilon_t \qquad t_a<t\leqslant t_b \tag{8-3-9}
$$

（8-3-7）在（t_a,t_b〕的模型是

$$
z_{i_t+d}=\beta_0+\beta_1z_{i_t+d-1}+\cdots+\beta_qz_{i_t+d-q}
+\sum_{j=1}^{s}\sum_{i=0}^{p_j}a_{ji}x_{j,i_t+d-i}+\varepsilon_t \qquad t_a<t\leqslant t_b \tag{8-3-10}
$$

式（8-3-8）、（8-3-9）、（8-3-10）都是简单的线性回归模型，用线性回归模型求解方法就可以解决。

关于门限值$\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_{r-1}$及$r$的选取，一般是通过经验分布确定。如我们取$\{x_{i,t}, 1 \leqslant t \leqslant N\}$作为门限变量，那么我们就要计算出它的经验概率密度，进而求出经验概率分布$F(x_{i,t})$，为了使每两个门限值之间出现的观察数据不至过少，一般选门限值为$\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_9$，使得$F(\bar{x}_1)=0.1$，$F(\bar{x}_2)=0.2, \cdots$，$F(\bar{x}_9)=0.9$。若从$\bar{x}_1, \cdots, \bar{x}_9$中挑选$r-1$个最优门限值，就是寻求使门限回归模型的AIC函数值为最小的那组个数为$r-1$的门限值。实际上一般取$r=2, 3$或4（即门限分点数为1，2或3）。对不同的r值再进行比较以确定最佳模型的门限值个数。

综上所述，建立门限回归、门限自回归模型的步骤可归纳如下：

（1）调用排序子程序ORDER，将门限变元$\{x_{i,t}, 1 \leqslant t \leqslant N\}$按照从小到大的顺序排列得到标号序列$\{i_t, 1 \leqslant t \leqslant N\}$。如果是自回归模型，则进行排序的就是观察序列本身。

（2）设m是给定的自回归或回归模型的最高阶数，p是当前拟合的模型阶数，对于模型式（8-3-5），模型阶数就是自变元的个数s，因此赋值方式为$m \Leftarrow p \Leftarrow s$（对照式(8-2-4)），从一阶模型着手拟合，即$p \Leftarrow 1$。

（3）对给定的一对门限分点$(t_a, t_b]$，调用子程序SUBLSM，对式（8-2-4）中的$a_{t,j}$，z_t赋值。

对于式（8-3-8），当$t_a < t \leqslant t_b$有

$$a_{t\,j} \Leftarrow x_{j,\,i_t+d} \qquad 1 \leqslant j \leqslant p$$

$$z_t \Leftarrow z_{i_t+d}$$

对于式（8-3-9），当$t_a < t \leqslant t_b$有

$$a_{t\,j} \Leftarrow z_{i_t+d-j} \qquad 1 \leqslant j \leqslant p$$

$z_t \Leftarrow z_{i_t+d}$

(4) 调用程序LSME，求与门限值 $(t_a, t_b]$ 相应的门限模型的参数、残差平方和及AIC值。

(5) 如果是模型（8-3-6）则令 $p \Leftarrow p+1$，若 $p<m$ 转(3)，直到选出 (t_a, t_b) 上AIC值最小的门限自回归模型；如果是模型（8-3-5）则转(6)。

(6) 对其它门限值重复(2)—(5)，将各段AIC值相加，得到总的AIC值。

(7) 对门限取值和分点数目的不同组合，重复(2)—(6)，比较不同模型的AIC值，得出最佳门限值及相应的门限回归（自回归）模型。

关于门限混合回归模型建立的步骤与前述大体相同，但情况更为复杂。在(3)中 a_{tj}，z_t 的计算公式为，对 $t_a<t\leqslant t_b$ 有

$$a_{tj} \Leftarrow \begin{cases} z_{i_t+d-j} & 1\leqslant j\leqslant m \\ x_{\nu, i_t+d-\mu} & j=\nu m+\mu,\ 1\leqslant \mu\leqslant m,\ 1\leqslant \nu\leqslant s \end{cases}$$

$z_t \Leftarrow z_t$

其中 m 是给定的最大迟后量，在(5)中对因变元及各自变元不超过 m 的迟后步数都要进行各种组合的搭配挑选，以得出各变量相应的最佳迟后步数，因此计算量是比较大的。我们可以在 $(t_a, t_b]$ 上采用8.2节中讲过的方法建立疏系数线性混合回归模型，即建立门限混合回归模型，这样可以大大减少计算量。

4.门限回归（自回归）模型的预报 设对观察序列 $\{z_t, 1\leqslant t\leqslant N\}$ 已建立了门限自回归模型

$$z_t=\beta_0^{(j)}+\sum_{i=1}^{p_j}\beta_i^{(j)} z_{t-i} \qquad 若 z_j<z_{t-d}\leqslant z_{j+1}$$

$$j=0,1,2,\cdots,r-1,r$$

门限值为 $z_1, z_2, \cdots, z_{r-1}$，并设 $z_0=-\infty$，$z_r=+\infty$，由第六章时间

序列的预报可知，平稳序列的m步最小方差预报应是预报量在已知观察值所构成的线性空间上的投影：

$$\hat{z}_t(m)=\hat{E}(z_{t+m}|z_1,z_2,\cdots,z_t)$$

当$m<d$时，z_{t+m-d}是已获得的量测数据，设对某个j，$z_{t+m-d}\in(z_j, z_{j+1}]$，则可以用相应区间上的$AR$模型进行预报

$$\hat{z}_t(m)=\beta_0^{(j)}+\beta_1^{(j)}\hat{z}_t(m-1)+\cdots+\beta_{m-1}^{(j)}\hat{z}_t(1)+ \\ +\beta_m^{(j)}z_t+\cdots+\beta_{p_j}^{(j)}z_{t+m-p_j} \quad (8\text{-}3\text{-}11)$$

这样，由实测数据和递推预报值$\hat{z}_t(1),\cdots,\hat{z}_t(m-1)$可以计算$m$步的预报值。

当$m>d$时，样本z_{t+m-d}尚未观测到，无法判断它属于哪个区间，也无法用上述方法进行m步最小方差预报，但可用近似预报代替，方法如下：

首先由前述方法求得精确的最小方差预报值$\hat{z}_t(j)$，$1\leqslant j\leqslant d$。求$\hat{z}_t(d+1)$时，由于$z_{t+m-d}=z_{t+1}$的值还没有观察到，因此用$\hat{z}_t(1)$近似，并由它判断选取哪个区间的模型，由量测值及已算出的预报值用式（8-3-11）计算$\hat{z}_t(d+1)$。依此类推，逐次计算$\hat{z}_t(d+2)$，$\hat{z}_t(d+3),\cdots,\hat{z}_t(m)$。由此便可得出任意$m$步的近似预报值。

我们把各步预报值用折线连接起来得到的曲线称为最终预报曲线。如果拟合模型稳定，那么该曲线将趋于一个常数或周期函数，对应于模型的差分方程，前者表征极限点，后者表征极限环。在8.3.5节中，我们将给出加拿大山猫捕获量的预测值，可以清楚地看到预报曲线是稳定的周期函数。

关于门限回归及门限混合回归的预报，其方法与门限自回归模型类似，但是当需要用到当前时刻t之后自变元的数据时，无法用近似值代替，而必须掌握自变元在时刻t之后的观察数据，才能进行外涨预报。

5.实例分析　在本节中将给出建立门限自回归、门限回归及

门限混合回归模型的实例。

例1 加拿大山猫捕获数量的门限自回归模型。

表8-9和图8-13为1821至1934年期间捕捉的加拿大山猫数的年度记录（以10为底取对数），共有114个观察值。这是时序分析学家做过很多研究的生态自然循环现象。在一个给定地区内的野兔和山猫的总数具有某种循环或振荡变化的型式，野兔是山猫的主要食物，当野兔数量很大时，山猫的数量增多，然而不久随着山猫成倍地增长，由于过量捕食的结果，野兔数量下降，而这又引起部分山猫饿死，从而山猫数量减少，这就完成个一个循环。如果野兔的数量比供山猫捕食所需野兔的数量大，则山猫数量又要增加（如果野兔数量总是不足，则会造成山猫绝种，但自然界并没有使它们的数量调整失去平衡），从观察数据可以看出其变化有近十年的明显周期，由于数据变化起伏很大，已事先对数据取对数以便建立门限自回归模型。

表8-9 1921—1934年加拿大山猫数据

269.	321.	585.	871.	1475.	2821.	3928.	5943.	4950.	2577.	523.	98.
184.	279.	409.	2285.	2685.	3409.	1824.	409.	151.	45.	68.	213.
546.	1033.	2129.	2536.	957.	361.	377.	225.	360.	731.	1638.	2725.
2871.	2119.	684.	299.	236.	245.	562.	1623.	3311.	6721.	4254.	687.
255.	473.	358.	784.	1594.	1676.	2251.	1426.	756.	299.	201.	229.
469.	736.	2042.	2811.	4431.	2511.	389.	73.	39.	49.	59.	188.
377.	1292.	4031.	3495.	537.	105.	153.	387.	758.	1307.	3465.	6991.
6313.	3794.	1836.	345.	382.	808.	1388.	2713.	3800.	3091.	2985.	3790.
674.	71.	80.	108.	229.	399.	1132.	2432.	3574.	2995.	1537.	529.
485.	662.	1000.	1520.	2657.	3396.						

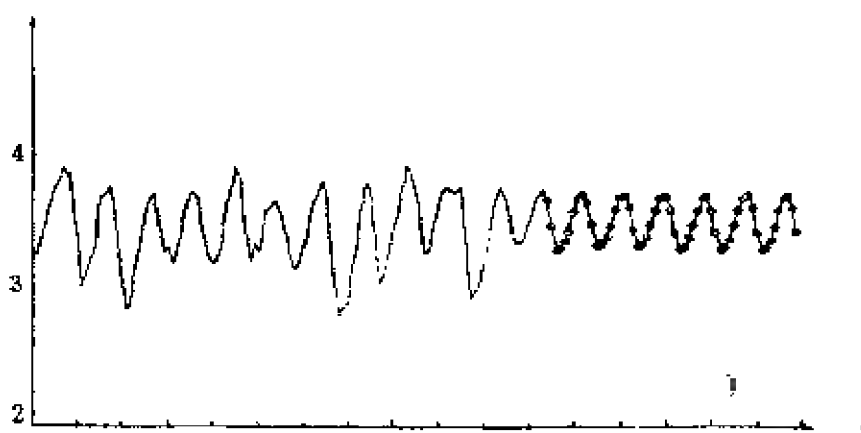


图8—13　山猫数据及预报曲线

门限值取为$z_1=3.1163$，门限迟后量$d=2$，得到的门限自回归模型是

$$
z_t=\begin{cases}
0.5239+1.0359z_{t-1}-0.1756z_{t-2}+0.1753z_{t-3} \\
-0.4339z_{t-4}+0.3457z_{t-5}-0.3032z_{t-6} \\
+0.2162z_{t-7}+0.0043z_{t-8}+\varepsilon_t^{(1)} \\
\qquad\qquad\qquad \text{当} z_{t-2}\leqslant 3.1163 \\
2.6559+1.4246z_{t-1}-1.1618z_{t-2}-0.1094z_{t-3} \\
+\varepsilon_t^{(2)} \qquad\qquad \text{当} z_{t-2}>3.1163
\end{cases}
$$

其中$\mathrm{Var}\ \varepsilon_t^{(1)}=0.0255$，$\mathrm{Var}\ \varepsilon_t^{(2)}=0.0516$

从图8–13中可以看到，最终预报曲线是一个周期 9 年的非对称函数，这表明该模型存在以 9 年为周期的极限环，上升段和下降段长度分别为 6 年和 3 年，它的极限环为（2.6226，2.8945，3.2525，3.4601，3.4257，3.2281，2.9793，2.7884，2.6639，…）

迟后量$d=2$表明山猫数量的增减与捕获量之间存在着迟后两年的依赖关系。

例2　流量、雨量门限模型的建立。本例对丹江口近十年的旬流量、雨量数据（见表 8—10）进行分析，建立门限回归、门

限自回归及门限混合回归模型，并对各种模型进行比较。我们用 $\{z_t\}$，$\{x_t\}$ 分别表示流量、雨量观察数据。众所周知，流量对雨量的依赖关系在少雨季节和多雨季节有明显差别，因此应当采用门限模型。门限值经过选择取为 $z_1=314.0$，$z_2=638.0$。以 σ_i^2表 (z_i, z_{i+1}) 上模型的残差方差 $(1\leqslant i\leqslant r,\ r=3)$，则该段上模型的AIC值为

$$AIC_i=\log(\sigma_i^2/N)+2p_i/N$$

其中 p_i 是该段模型自回归阶数。本例中样本个数$N=391$。整个门限自回归模型的AIC值等于各段AIC函数值之和

$$AIC=\sum_{i=1}^{r}AIC_i$$

取门限迟后$d=0$，1进行流量自回归模型，流量、雨量门限回归及门限混合回归模型的拟合。结果表明，$d=0$时拟合效果较好，这说明流量的变化在本旬之内就出现影响。拟合结果见表8-11，这些结果是使用程序TREG进行计算的。

比较表8—11中各模型的拟合残差及AIC值可以看出，模型(1)比(2)好，(4)比(5)好，这说明雨量对流量的影响很大，它超过了流量自身历史状况的影响，因此分析流量的变化规律时，雨量的作用是不可忽视的。另外，由表8—11中还可以看出，门限模型(4)—(7)都比非门限模型(1)—(3)好，这是由于流量的变化幅度很大，不同取值范围流量的均值、方差相差悬殊，因此采用门限模型是合理的。表8—11所列举的模型中，(7)拟合效果最好，该门限混合回归模型既考虑了雨量对流量的作用，又考虑了流量自身的历史数据，而且采用了疏系数模型。由于特别关心的是洪水预测问题，因此希望对于大流量的拟合残差应尽量小。在模型(7)中拟合残差均方 $\sigma_3=682.3$，而该段流量数据本身的均方 $\sigma_3^0=1367.8$，二者之比 $\sigma_3/\sigma_3^0=682.3/1367.8=0.4988$，这在表8—13中是拟合效果最好的。

表8—10 丹江口旬流量，雨量数据（1流量，2雨量）

1	542.	330.	347.	695.	397.	319.	372.	489.	963.	854.	3871.	1369.	890.	888.	330.
2	2.	10.	1.	9.	1.	5.	2.	20.	15.	35.	51.	15.	25.	15.	20.
1	282.	223.	1220.	531.	590.	304.	223.	362.	2513.	1051.	4060.	1430.	583.	1830.	881.
2	8.	10.	71.	19.	40.	22.	25.	15.	105.	1.	115.	4.	27.	29.	15.
1	783.	594.	427.	333.	389.	291.	251.	172.	210.	217.	230.	301.	402.	613.	1062.
2	18.	3.	1.	0.	2.	1.	0.	2.	2.	2.	10.	8.	9.	23.	18.
1	1388.	1551.	1703.	834.	978.	3000.	1319.	840.	684.	397.	1742.	1920.	683.	380.	709.
2	16.	59.	10.	14.	68.	42.	46.	12.	30.	10.	98.	38.	17.	34.	32.
1	847.	6429.	1587.	2881.	2001.	844.	521.	511.	457.	398.	386.	290.	798.	221.	199.
2	30.	107.	40.	19.	33.	0.	0.	18.	1.	5.	3.	0.	9.	2.	2.
1	244.	314.	306.	284.	576.	603.	1340.	1159.	2267.	883.	590.	948.	5395.	1585.	1172.
2	4.	7.	3.	4.	18.	17.	28.	19.	43.	31.	7.	85.	59.	49.	64.
1	2086.	764.	726.	638.	2480.	1300.	1067.	920.	4279.	712.	1720.	2245.	1858.	735.	518.
2	30.	19.	33.	32.	67.	48.	25.	33.	54.	40.	10.	52.	6.	1.	0.
1	433.	357.	341.	353.	285.	352.	287.	298.	277.	965.	1232.	383.	761.	1533.	786.
2	0.	4.	3.	0.	10.	7.	4.	1.	22.	17.	31.	12.	26.	22.	28.
1	507.	1589.	919.	310.	1545.	3780.	3710.	1190.	430.	346.	282.	3330.	1070.	485.	455.
2	6.	64.	3.	25.	62.	120.	17.	29.	15.	11.	46.	79.	7.	10.	14.
1	661.	555.	486.	1394.	463.	887.	283.	192.	160.	215.	199.	225.	208.	341.	257.
2	24.	4.	14.	19.	2.	0.	0.	4.	1.	3.	7.	6.	1.	0.	14.

续上表

1	322.	310.	691.	1538.	1086.	2677.	1207.	675.	641.	2189.	1211.	1865.	1580.	1378.	2700.
2	6.	21.	48.	13.	53.	40.	13.	33.	22.	58.	12.	66.	59.	82.	18.
1	691.	522.	3488.	1408.	2165.	4342.	674.	792.	694.	514.	382.	339.	148.	188.	261.
2	24.	32.	125.	9.	65.	79.	17.	13.	10.	0.	1.	1.	3.	0.	5.
1	281.	208.	284.	245.	431.	160.	784.	333.	658.	751.	540.	1607.	1840.	3970.	433.
2	5.	3.	6.	2.	3.	23.	5.	5.	8.	22.	5.	34.	81.	33.	36.
1	686.	487.	2450.	922.	589.	3304.	2077.	975.	1526.	3314.	1354.	8240.	1309.	1200.	862.
2	23.	26.	65.	24.	47.	123.	32.	20.	91.	122.	30.	91.	15.	18.	10.
1	721.	550.	635.	385.	1553.	361.	213.	182.	301.	271.	446.	315.	352.	263.	248.
2	14.	8.	11.	2.	8.	0.	2.	1.	8.	0.	1.	5.	5.	5.	3.
1	404.	1697.	1390.	1012.	1039.	630.	832.	1875.	3430.	3610.	970.	2150.	1580.	570.	1710.
2	36.	60.	10.	48.	9.	35.	15.	55.	126.	9.	26.	57.	13.	25.	81.
1	1970.	5780.	9990.	3200.	1410.	846.	1228.	682.	532.	477.	263.	577.	430.	337.	310.
2	60.	141.	77.	25.	22.	14.	6.	0.	12.	0.	0.	0.	1.	1.	1.
1	578.	647.	519.	483.	584.	442.	675.	913.	1342.	893.	1093.	2235.	838.	1147.	790.
2	24.	6.	6.	7.	6.	11.	17.	39.	19.	19.	72.	8.	31.	38.	16.
1	900.	1200.	636.	354.	3530.	2070	930.	670.	750.	2000.	800.	734.	816.	458.	445
2	46.	26.	15.	2.	139.	25.	19.	5.	64.	23.	18.	26.	3.	2.	7
1	339.	191.	303.	238.	313.	263.	233.	226.	357.	362.	631.	667.	1065.	1487.	1170
2	0.	1.	1.	2.	8.	2.	0.	0.	1.	18.	23.	10.	34.	54.	25.

续上表

1	1636.	945.	605.	522.	871.	785.	4100.	1930.	1700.	1720.	722.	614.	534.	404.	345.
2	35.	28.	11.	12.	43.	104.	56.	63.	35.	33.	23.	12.	23.	2.	18.
1	301.	316.	1088.	638.	477.	399.	412.	370.	346.	208.	202.	212.	211.	206.	209.
2	2.	67.	27.	4.	0.	2.	16.	0.	0.	1.	1.	4.	1.	2.	12.
1	309.	242.	537.	377.	422.	574.	676.	546.	1287.	696.	790.	5840.	2100.	2180.	1180.
2	7.	4.	3.	17.	4.	51.	8.	67.	13.	27.	40.	144.	53.	64.	16.
1	864.	570.	735.	1050.	561.	343.	430.	373.	229.	527.	346.	249.	211.	160.	168.
2	8.	26.	41.	30.	4.	24.	8.	10.	26.	9.	3.	1.	2.	1.	0.
1	199.	133.	143.	202.	220.	381.	342.	286.	255.	484.	725.	932.	1370.	635.	926.
2	7.	[illegible]	1.	4.	20.	5.	5.	15.	8.	15.	41.	23.	39.	5.	47.
1	514.	578.	264.	5789.	1782.	2293.	1230.	937.	1340.	4838.	5296.	1014.	552.	430.	341.
2	30.	28.	81.	137.	35.	41.	31.	57.	18.	93.	67.	1.	15.	10.	9.
1	311.														
2	14.														

表8—11　流量、雨量拟合结果

	模型名称	拟合模型	残差	AIC值
(1)	流量—雨量一元回归模型	$z_t=320.760+29.077x_t+\varepsilon_t$	$\hat\sigma_1=439.80$ $\hat\sigma_2=426.54$ $\hat\sigma_3=1111.00$	38.37
(2)	流量自回归模型	$z_t=588.434+0.340z_{t-1}+0.107z_{t-2}+\varepsilon_t$	$\hat\sigma_1=470.80$ $\hat\sigma_2=444.78$ $\hat\sigma_3=1432.80$	39.14
(3)	流量雨量多元回归模型	$z_t=16.998+23.554x_t+15.172x_{t-1}$ $+2.852x_{t-2}+\varepsilon_t$	$\hat\sigma_1=348.74$ $\hat\sigma_2=394.77$ $\hat\sigma_3=954.33$	37.54
(4)	流量雨量门限回归模型	$z_t=235.3+0.81x_t+\varepsilon_t^{(1)}$ $z_t=430.6+2.257x_t+\varepsilon_t^{(2)}$ $z_t=666.0+26.46x_t+\varepsilon_t^{(3)}$	$\hat\sigma_1=47.12$ $\hat\sigma_2=92.12$ $\hat\sigma_2=1080.49$	30.77

续上表

（5）	流量门限自回归模型	$z_t=209.9+0.1047z_{t-1}+\varepsilon_t^{(1)}$ $z_t=407.1+0.0410z_{t-1}+0.0311z_{t-2}+\varepsilon_t^{(2)}$ $z_t=1339.4+0.1275z_{t-1}+0.0316z_{t-2}-$ $-6.0484z_{t-3}-0.1115z_{t-4}-0.0565z_{t-4}$ $+0.3528z_{t-6}+\varepsilon_t^{(3)}$	$\sigma_1=45.65$ $\sigma_2=89.97$ $\sigma_3=1309.48$	31.11
（6）	流量雨量门限混合回归模型	$z_t=231.238+1.750x_{t-1}+0.723x_{t-6}-$ $-1.454x_{t-7}+0.386x_{t10}+\varepsilon_t^{(1)}$ $z_t=723.618-4.137x_t+1.417x_{t-2}$ $+1.083x_{t-3}-1.036x_{t-4}-0.880x_{t-5}$ $+0.363x_{t-6}+0.901x_{t-7}-0.248x_{t-8}$ $-1.037x_{t-9}-0.409x_{t-10}-1.313x_{t12}$ $-1.442x_{t-13}-1.956x_{t-14}+\varepsilon_t^{(2)}$ $z_t=-260.012+26.425x_t+15.839x_{t-1}$ $13.037x_{t-14}+\varepsilon_t^{(3)}$	$\sigma_1=20.05$ $\sigma_2=10.659$ $\sigma_3=711.80$	26.55

续上表

（7）	流量—流量雨量门限混合回归模型	$z_t=206.176+0.089z_{t-1}-0.046z_{t-3}+0.035z_{t-4}+0.03z_{t-8}+0.014z_{t-9}+1.524x_{t-3}-1.424x_{t-6}-0.914x_{t-8}-0.494x_{t-9}+\varepsilon_t^{(1)}$	$\hat{\sigma}_1=17.82$	25.412
		$z_t=597.219-0.136z_{t-1}-0.021z_{t-4}+0.039z_{t-5}+0.036z_{t-6}-0.039z_{t-8}-0.018z_{t-10}-2.398x_t+1.739x_{t-1}+2.769x_{t-2}+1.26x_{t-3}-1.279x_{t-5}-2.051x_{t-6}-0.754x_{t-7}+\varepsilon_t^{(2)}$	$\hat{\sigma}_2=32.55$	
		$z_t=-140.525+26.096x_t+15.126x_{t-1}+8.76x_{t-11}+\varepsilon_t^{(3)}$	$\hat{\sigma}_3=682.32$	

注：(4)～(7)诸门限模型的门限值均取为$z_1=3140.0$　$z_2=638.0$

应当指出，对观察数据采取开平方或取对数等变换处理，可能进一步改善拟合效果，这里不作进一步讨论。

8.4. 双线性模型和指数自回归模型

本节介绍两种非线性模型：双线性模型，指数自回归模型，每种模型都给出实例及参数估计方法。

1.双线性模型 在时间域内处理非线性模型的主要困难在于迄今还没有找到能用有限个参数表示的一般表达式。近代控制理论中提出的一类"双线性模型"开拓了很好的可能性，由于只包含有限个参数，使得用这种模型进行拟合具有实现的可能，因而该模型可看作是ARMA类型的模型在非线性情形下的自然推广。

双线性模型的标量形式可表为

$$x_t+a_1x_{t-1}+\cdots+a_nx_{t-n}=e_t+b_1e_{t-1}+\cdots+$$

$$+b_me_{t-m}+\sum_{i=1}^{k}\sum_{j=1}^{l}c_{ij}e_{t-i}x_{t-j} \qquad (8\text{-}4\text{-}1)$$

显然，若对所有i，j有$c_{ij}=0$，则式(8-4-1)就是ARMA模型。

虽然双线性模型在控制理论中已有相当深入的探讨，这类模型的统计分析却还不成熟。有些学者对某些特定结构的双线性时间序列模型做过讨论，但它的一般性质还有待进一步研究。

如果双线性模型的阶数(即总的参数数目)事先已定，则其参数估计原则上与线性模型没有区别。设e_t是正态过程，模型(8-4-1)在$t=1,2,\cdots,N$时的观察值x_t为已知，当N很大时近似有似然函数

$$L(\theta)\cong\exp\left(-\frac{1}{2\sigma_e^2}\sum_{t=2}^{N}e_t^2\right) \qquad (8\text{-}4\text{-}2)$$

其中θ为参数集合，即$\theta=\{a_i, b_j, c_{rs}; i=1,\cdots,n, j=1,\cdots, m, r=1,\cdots,k, s=1,\cdots,l\}$，对每一个$\theta$值，$\varepsilon_t$可以由式（8-4-1）递推计算。所以$\theta$的极大似然估计可将

$$V(\theta)=\sum_{t=1}^{N}\varepsilon_t^2 \tag{8-4-3}$$

对θ的每个元素取极小化而获得。首先定出一组初始估计$\hat{\theta}_0$，然后用递推算法来求使$V(\theta)$为最小的$\hat{\theta}$。索巴·雷欧(*Subba Rao*)采用牛顿—拉夫森方法进行递推，并提出了另一种“反复残差”的方法。譬如有一阶双线性模型

$$x_{t+1}=ax_t+bx_t\varepsilon_t+\varepsilon_{t+1} \tag{8-4-4}$$

将它写成

$$(1+bx_tB)\varepsilon_{t+1}=x_{t+1}-ax_t \tag{8-4-5}$$

对于很小的b，近似有

$$\begin{aligned}\varepsilon_{t+1}&=(1-bx_tB)(x_{t+1}-ax_t)\\&=x_{t+1}-ax_t+bx_t^2+abx_tx_{t-1}\end{aligned} \tag{8-4-6}$$

忽略乘积项x_tx_{t-1}，可以用线性最小二乘法求a和b的估计（若不忽略乘积项，则需用非线性最小二乘法估计a和b)，我们把a，b的初始估计记作$\hat{a}_0$，$\hat{b}_0$，一旦求出$\hat{a}_0$，$\hat{b}_0$，便可由下式递推计算求出$\hat{\varepsilon}_t$（从$\hat{\varepsilon}_1$开始）

$$\hat{\varepsilon}_{t+1}=x_{t+1}-\hat{a}_0x_t-\hat{b}_0x_t\hat{\varepsilon}_t,\ t=1,\cdots,N \tag{8-4-7}$$

而给出$\{x_t, \hat{\varepsilon}_t; t=1,\cdots,N\}$，然后由极小化

$$\sum_{t=1}^{N}(x_{t+1}-ax_t-bx_t\hat{\varepsilon}_t)^2 \tag{8-4-8}$$

来估计新的a和b，以上步骤反复进行直到收敛为止。索巴·雷欧研究如下一阶双线性模型

$$x_{t+1}=0.4x_t+0.4x_t\varepsilon_t+\varepsilon_{t+1}$$

其中ε_t为独立$N(0,1)$变量，根据200个观察值求得估计为

(1) 牛顿一拉夫森方法：$\hat{a}=0.4978$；$\hat{b}=0.3935$

(2) 反复残差法：　　　$\hat{a}=0.5135$；$\hat{b}=0.4188$

下面再来看双线性模型在渥尔夫（*Wolf*）太阳黑子序列分析中的应用。在1927年耶尔（*Yule*）最早用自回归模型研究这一序列，太阳黑子数据1700年开始便有纪录可查（在本书表8—6中列出），它的主要特点是其活动"周期"在9到14年左右的范围内变化，平均周期约为11年，这一特点促使耶尔用 *AR*(2) 模型拟合。后来还有一些学者用过AR(3)、*AR*(5)模型进行拟合。然而太阳黑子数据中具有的另一特点是它的"上升"和"下降"坡度不同，即在每一循环周期中，升到最高点比降到最低点更快些，这一性质表明模型可能应是非线性的。为此，哈根（*Haggan*）和索巴·雷欧既用线性也用双线性模型做拟合，并比较所得的结果。他们用246个观察数据（均值为43.53），并在拟合线性模型时将均值扣除，线性模型的阶数按AIC准则确定。下面归纳几种模型的拟合结果。

线性模型拟合结果：

(1) *AR*(*n*)模型 $x_t+a_1x_{t-1}+\cdots+a_nx_{t-n}=\varepsilon_t$。最佳模型为AR(9)，参数是 $\hat{a}_1=-1.21$，$\hat{a}_2=0.47$，$\hat{a}_3=0.12$，$\hat{a}_4=-0.14$，$\hat{a}_5=0.16$，$\hat{a}_6=-0.09$，$\hat{a}_7=0.08$，$\hat{a}_8=-0.09$，$\hat{a}_9=-0.10$。估计残差方差$\hat{\sigma}^2=194.43$，AIC=1316.44。

(2) 疏系数*AR*模型．最佳疏系数*AR*模型包含三项(1,2,9)系数为$\hat{a}_1=-1.24$，$\hat{a}_2=0.54$，$\hat{a}_3=-0.15$，$\hat{\sigma}^2=199.2$，AIC=1310.39。

(3) ARMA模型 $x_t+a_1x_{t-1}+\cdots+a_nx_{t-n}=\varepsilon_t+b_1\varepsilon_{t-1}+\cdots+b_m\varepsilon_{t-m}$。最佳模型为ARMA(6,6)，估计参数为$\hat{a}_1=-0.52$，$\hat{a}_2=-0.44$，$\hat{a}_3=-0.51$，$\hat{a}_4=1.09$，$\hat{a}_5=0.07$，$\hat{a}_6=-0.25$，$\hat{b}_1=0.71$，$\hat{b}_2=-0.07$，$\hat{b}_3=-1.09$，$\hat{b}_4=-0.08$，$\hat{b}_5=0.41$，$\hat{b}_6=0.42$。残差方差$\hat{\sigma}^2=185.27$，AIC=1309.8

双线性模型拟合结果：在式（8-4-1）中取 $n=l$，并取消该式右边的纯MA项，即

$$x_t+a_1x_{t-1}+\cdots+a_nx_{t-n}=$$

$$d+\sum_{i=1}^{k}\sum_{j=1}^{n}c_{ij}e_{t-i}x_{t-j}$$

式中包含有常数 d 以便于拟合不扣除均值的原始数据。利用AIC准则选得的最佳模型是 $n=3$，$k=4$。估计的残差方差 $\hat{\sigma}^2=149.60$，AIC=1243.93。模型参数为 $\hat{d}=10.9132$，$\hat{a}_1=-1.93$，$\hat{a}_2=1.46$，$\hat{a}_3=-0.27$，$\hat{c}_{ij}$ 按矩阵元素排列如下：

$$\begin{matrix}-0.55\times10^{-2} & -0.57\times10^{-2} & -0.17\times10^{-2}\\ 0.32\times10^{-2} & -0.56\times10^{-2} & 0.71\times10^{-2}\\ -0.18\times10^{-2} & -0.82\times10^{-2} & 1.45\times10^{-2}\\ 0.08\times10^{-2} & 0.58\times10^{-2} & -0.03\times10^{-2}\end{matrix}$$

由前述可以看出，对太阳黑子数据进行拟合，用双线性模型比用线性模型有十分明显的改进。

2.指数自回归模型 指数自回归模型是尾崎（*Ozaki*）和哈根（*Haggan*）提出的，用以复现非线性随机振动的某些特性。我们知道，非线性随机振动通常可以用以下二阶微分方程描述：

$$\ddot{x}(t)+f\{\dot{x}(t)\}+g\{x(t)\}=y(t) \qquad (8\text{-}4\text{-}9)$$

其中"阻尼力" $f\{\cdot\}$ 和"恢复力" $g\{\cdot\}$ 均为非线性函数，但 $y(t)$ 为随机"输入"。杜芬（*Duffing*）方程

$$\ddot{x}(t)+c\dot{x}(t)+ax(t)+b\{x(t)\}^3=y(t)$$

和温德波（*Van der Pol*）方程

$$\ddot{x}(t)+f\{\dot{x}(t)\}+ax(t)=y(t)$$

就是式（8-4-9）的两个例子。如果 $f\{\cdot\}$、$g\{\cdot\}$ 都是线性函数，则式（8-4-9）就是熟知的二阶线性微分方程

$$\ddot{x}(t)+a_1\dot{x}(t)+a_2x(t)=y(t)$$

当$y(t)$为白噪声时，$x(t)$为连续$AR(2)$过程，$x(t)$的功率谱密度在某一频率f_1处出现峰值。亦即系统在白噪声输入下，在输出$x(t)$中将有近似于周期f_1的过程，这个f_1值只和α_1，α_2有关。如果输入是周期的，如$y(t)=A\cos 2\pi f_0 t$，则输出将是

$$x(t)=A'\cos(2\pi f_0 t+\beta)$$

其振幅$A'=A/|-\omega_0^2+j\alpha_1\omega_0+\alpha_2|$，$\omega_0=2\pi f_0$。即$x(t)$的振幅依输入频率而连续变化。

上面所说的是众所周知的线性系统情形。如果是非线性系统，则当输入为白噪声时，输出可能还是振荡的，但近似的周期f_1却可能随振荡的振幅变化，即f_1值取决于$x(t)$的大小，这就是所谓“依赖于振幅的频率”。此外，当输入是频率为f_0的周期量时，非线性系统可能具有近似于周期变化的输出

$$x(t)\cong B\cos(2\pi f_0 t+\beta)$$

但振幅B可能是f_0的不连续函数，即当f_0变到某一值时，B会从一个值“跳”到另一个值。这种情况在解杜芬方程时可以看出，我们称之为“突跳”现象。

尾崎提出了产生“依赖于振幅的频率”这一效应的离散参数时间序列模型为下面$AR(2)$模型

$$x_t-a_1x_{t-1}-a_2x_{t-2}=\varepsilon_t \tag{8-4-10}$$

这里的a_1，a_2可依赖于x_{t-1}。这种依赖关系为指数函数形式，即

$$a_1=\alpha_1+\beta_1\exp(-\gamma x_{t-1}^2),\quad a_2=\alpha_2+\beta_2\exp(-\gamma x_{t-1}^2) \tag{8-4-11}$$

如果我们还把式（8-4-10）看作是“线性”的，那么该模型的“共振频率”将是使$|1-a_1e^{-j2\pi f}-a_2e^{-j4\pi f}|$为最小的频率，从而将随$x_{t-1}$的大小变化（当$x_{t-1}$很大时，$a_1\cong\alpha_1$，$a_2\cong\alpha_2$，而当$x_{t-1}$很小时，$a_1\cong\alpha_1+\beta_1$，$a_2\cong\alpha_2+\beta_2$，其间变化关系是连续的）。这个模型也会产生“突跳”现象和极限环。按照式（8-4-11）选择a_1，a_2，则式（8-4-10）称为二阶指数自回归模型。一

般情况下的n阶指数自回归模型表达式为

$$x_t=[\alpha_1+\beta_1\exp(-\gamma x_{t-1}^2)]x_{t-1}+\cdots+[\alpha_n+ \\ +\beta_n\exp(-\gamma x_{t-n}^2)]x_{t-n}+\varepsilon_t \qquad (8\text{-}4\text{-}12)$$

它的参数估计方法为

(1) 先固定γ值，按线性最小二乘方法做x_t对x_{t-1}, $x_{t-1}\exp(-\gamma x_{t-1}^2)$,…的回归分析估计$\hat{\alpha}_1$，$\hat{\beta}_1$，…$\hat{\alpha}_n$，$\hat{\beta}_n$，用AIC准则确定阶数$n$。

(2) 在一定范围内取不同的γ值重复(1)的做法，按AIC准则选择最适宜的γ值。通常γ取为能使$\exp(-\gamma x_{t-1}^2)$在(0,1)范围内为宜。

哈根和尾崎用式 (8-4-12) 拟合了经对数换算和扣除均值后的加拿大山猫数据（1821年—1934年），依照上述算法求得最好的模型是：$n=11$，$\gamma=3.89$，而α_i和β_i的估计值如下

i	1	2	3	4	5	6	7	8	9	10	11
$\hat{\alpha}_i$	1.09	−0.28	0.27	−0.45	0.41	−0.36	0.22	−0.10	0.22	−0.07	−0.38
$\hat{\beta}_i$	0.01	−0.49	−0.06	0.30	−0.54	0.61	−0.53	0.30	−0.18	0.18	0.16

估计的残差方差$\hat{\sigma}_\varepsilon^2=0.0321$。而坎贝尔（*Campbell*）和瓦克尔（*Walker*）用$AR(2)$叠加一个简谐分量的模型进行拟合，残差方差是0.0437；汤家豪等所拟合的门限自回归模型（根据$x_{t-2}\leqslant 3.1163$和$x_{t-2}>3.1163$，模型阶数分别为8和3），残差方差是0.0360。用指数自回归模型进行拟合残差方差比后二者来得小。上面求得的指数自回归模型具有近似为十年的极限环。x_{t-1}较小时周期约为10.1年，x_{t-1}较大时周期约为9.4年，这和实际数据相当符合。

8.5. 自适应AR模型

1. 自适应 *AR* 模型的构造原理 这里的自适应模型是指在某种程度上能够根据量测数据（输入）和估计结果（输出）自行调整模型参数，并随着新数据的到来，通过递推算法自动地对系统参数加以修正，使其接近于某种最佳值。

这种模型的建立可以看成由以下几步构成：(1)对现场实时测量的时间过程进行采样，或事先记录的时间序列数据按先后顺序及一定时间间隔输入。(2)对一定长度的相邻输入数据给予不同的加权并求其加权和。(3)自动调整加权系数使得到的模型为最好。图8—15表示这种自适应模型的构造原理，其中x_t为当前

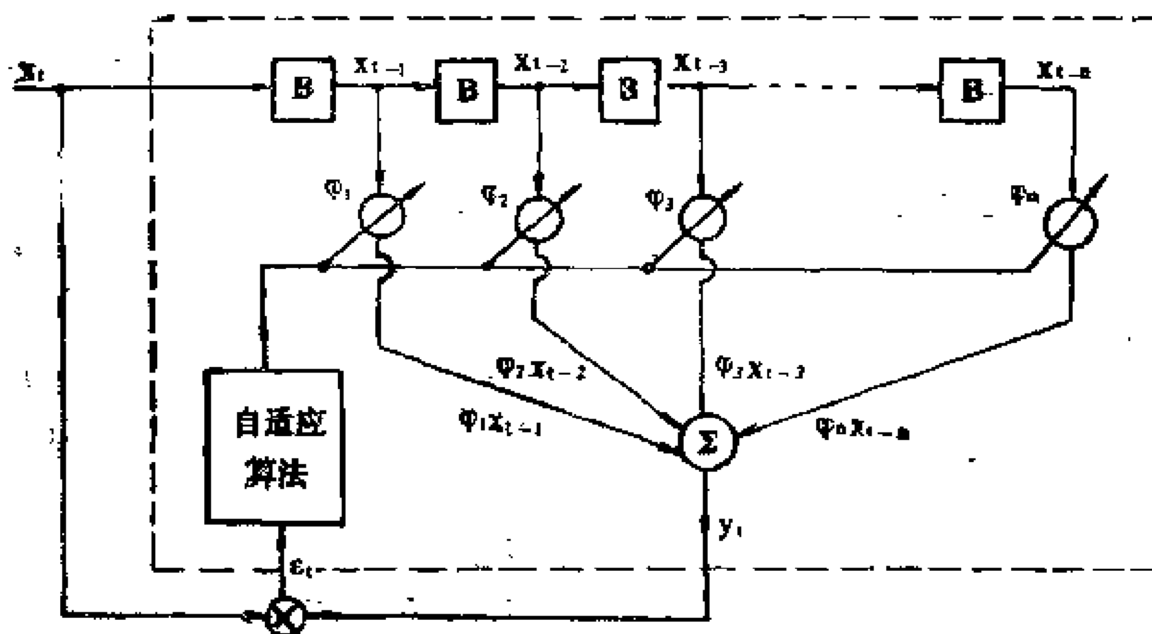


图8-14　自适应AR模型构造原理

的输入数据，$\varphi, \cdots, \varphi_n$是对既往$n$个数据的加权系数，$y_t$是加权和：

$$y_t = \varphi_1 x_{t-1} + \cdots + \varphi_n x_{t-n} \qquad (8\text{-}5\text{-}1)$$

如果参数φ_i是根据使$\varepsilon_t = x_t - y_t$的方差最小来确定，则方程 式

$$x_t - \varphi_1 x_{t-1} - \cdots - \varphi_n x_{t-n} = \varepsilon_t \tag{8-5-2}$$

相当于最小方差拟合的$AR(n)$模型，ε_t为拟合残差。下面将要证明，如果需要建模的数据在统计上是平稳的，则残差的均方值是参数的二次函数，用最速下降法可以自动地实现均方误差的极小化。设在k时刻（或第k次递推计算）的自回归参数，用向量表示为

$$\boldsymbol{\varphi}_k = [\varphi_{1k}, \varphi_{2k}, \cdots, \varphi_{nk}]^T \tag{8-5-3}$$

而被加权的输入量为

$$X_k = [x_{k-1}, x_{k-2}, \cdots, x_{k-n}]^T \tag{8-5-4}$$

k时刻的加权输出y_k和残差ε_k分别为

$$y_k = \sum_{i=1}^{n} \varphi_{ik} x_{k-i} = \boldsymbol{\varphi}_k^T X_k = X_k^T \boldsymbol{\varphi}_k \tag{8-5-5}$$

$$\varepsilon_k = x_k - y_k = x_k - \boldsymbol{\varphi}_k^T X_k \tag{8-5-6}$$

残差的平方及均方值分区为

$$\varepsilon_k^2 = x_1^2 - 2x_k X_k^T \boldsymbol{\varphi}_k + \boldsymbol{\varphi}_k^T X_k X_k^T \boldsymbol{\varphi}_k \tag{8-5-7}$$

$$E[\varepsilon_k^2] = E[x_k^2] - 2R_{xX}^T \boldsymbol{\varphi}_k + \boldsymbol{\varphi}_k^T R_{XX} \boldsymbol{\varphi}_k \tag{8-5-8}$$

其中

$$R_{xX} = E[x_k X_k] = E\begin{pmatrix} x_{k-1}x_k \\ x_{k-2}x_k \\ \vdots \\ x_{k-n}x_k \end{pmatrix} \tag{8-5-9}$$

$$R_{XX} = E[X_k X_k^T] = E\begin{pmatrix} x_{k-1}x_{k-1} & x_{k-1}x_{k-2} \cdots x_{k-1}x_{k-n} \\ x_{k-2}x_{k-1} & x_{k-2}x_{k-2} \cdots x_{k-2}x_{k-n} \\ \vdots & \\ x_{k-n}x_{k-1} & x_{k-n}x_{k-2} \cdots x_{k-n}x_{k-n} \end{pmatrix} \tag{8-5-10}$$

由式（8-5-8）可见，对平稳过程而言，均方误差是参数$\boldsymbol{\varphi}$的二次

函数，这样的性能函数几何图形如同碗状，自适应过程就是自动修正参数以确定碗底。在非平稳情况下，这个碗的形状和碗底的位置可以是变动的。这时自适应过程则需跟踪碗底。做为本节的重点，是讨论对未知平稳过程做自适应建模的过渡过程，以及当自适应过程结束后的稳态情形。

将式（8-5-8）的均方误差对参数向量求导，可得性能函数在任一点的梯度为：

$$\nabla[E\varepsilon_k^2]=-2R_{xX}+2R_{XX}\varphi_k \tag{8-5-11}$$

令梯度为零，可得使均方误差为最小的最优解

$$\varphi_{LMS}=R_{XX}^{-1}R_{xX} \tag{8-5-12}$$

此时的最小均方误差为

$$\varepsilon_{min}^2\triangleq E[\varepsilon_k^2]|_{min}=E[x_k^2]-\varphi_{LMS}^T R_{xX} \tag{8-5-13}$$

如果将式（8-5-13）代入式（8-5-8）可得 $E[\varepsilon_k^2]$ 的另一种表达式

$$E[\varepsilon_k^2]=\varepsilon_{min}^2+(\varphi_k-\varphi_{LMS})^T R_{XX}(\varphi_k-\varphi_{LMS}) \tag{8-5-14}$$

2.最速下降法及其渐近性质 用最速下降法寻找使均方误差达最小的参数，通常是从参数的一组初估计值开始，根据与这组初始值所对应的梯度向量做下组估计值。参数向量的变化是沿着梯度向量的相反方向。如果参数向量的变化使均方误差变化的趋势是减小的，则搜索过程将收敛到均方误差的最小点。

最速下降法的关系式为

$$\begin{aligned}\varphi_{k+1}&=\varphi_k-\mu\nabla[E\varepsilon_k^2]\\&=\varphi_k-2\mu R_{XX}\varphi_k+2\mu R_{xX}\end{aligned} \tag{8-5-15}$$

这里φ_{k+1}的计算没有矩阵求逆。

将式（8-5-14）对φ_k求导并令V_k表φ_k和φ_{LMS}之差

$$V_k=\varphi_k-\varphi_{LMS} \tag{8-5-16}$$

可得梯度的另一种表达式

$$\nabla=2R_{XX}V_k \tag{8-5-17}$$

为分析最速下降法自适应过程的性质，把式（8-5-15）画成带有反馈的信号流通图比较方便（见图8—15），它类似于一个一阶n维随动系统，V_k是使系统动作的“误差信号”。

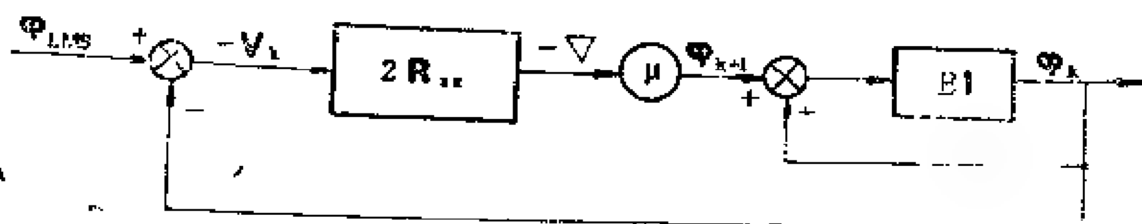


图6-15　带反馈的信号流通图

由于x_t通常是相关的，即矩阵R_{xx}的非主对角线元素一般非零，因此上述n维系统是互相关连（或称耦合）的。但考虑到R_{xx}是对称正定阵，可以将其正交分解得

$$R_{xx}=Q^{-1}\Lambda Q \tag{8-5-18}$$

其中Λ是特征值对角阵，即$\Lambda=\mathrm{diag}[\lambda_1,\lambda_2,\cdots,\lambda_n]$ Q是模态矩阵，且$Q^{-1}=Q^T$。利用这种正交分解可将式（8-5-15）写成

$$\varphi_{k+1}=\varphi_k-2\mu(2Q\Lambda Q^{-1}V_k) \tag{8-5-19}$$

令

$$\left.\begin{aligned}V_k'&=Q^{-1}V_k\\ \varphi_k'&=Q^{-1}\varphi_k\\ \varphi_{LMk}'&=Q^{-1}\varphi_{LMS}\end{aligned}\right\} \tag{8-5-20}$$

可得

$$V_{k+1}'-V_k'+2\mu\Lambda V_k'=0 \tag{8-5-21}$$

或

$$V_{k+1}'-(I-2\mu\Lambda)V_k'=0 \tag{8-5-22}$$

这是解耦后的参数向量误差方程式，它是齐次向量差分方程，根据它可以画出无耦合的n维系统信号流通图（见图8—16）。若初始条件为$V_0'=\varphi_0'-\varphi_{LMS}'$，则式（8-5-22）的解为

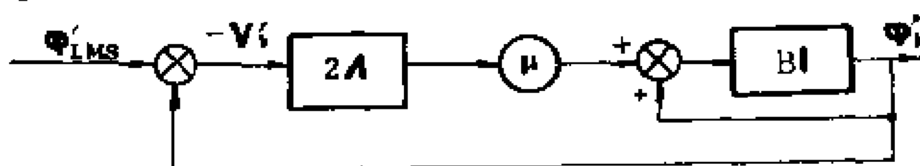


图8-16　无耦合n维系统信号流通图

$$V'_k=(I-2\mu\Lambda)^k V'_0$$

对于第i个系统（i=1,2,…,n），V'_k的第i个元素为

$$v'_{ik}=(1-\mu\lambda_i)^k v'_0=r_i^k v'_0 \tag{8-5-24}$$

其中

$$r_i=1-2\mu\lambda_i \tag{8-5-25}$$

系统渐近稳定或v'_{ik}收敛到零的条件为

$$|r_i|<1$$

或

$$0<\mu<1/\lambda_i$$

要使n个v'_{ik}都收敛到零，则μ的选择应满足

$$0<\mu<1/\lambda_{max} \tag{8-5-26}$$

λ_{max}是R_{XX}矩阵的最大特征值，在条件（8-5-26）满足的情况下

$$\varphi_\infty=\varphi_{LMk} \tag{8-5-27}$$

μ的选择不但决定了φ_k是否收敛，而且也影响收敛速度，即自适应过程进行的快慢。

如果用一个指数函数e^{-1/τ_i}表示r_i，或

$$\log r_i=-1/\tau_i \tag{8-5-28}$$

近似有

$$\tau_i\cong\frac{1}{2\mu\lambda_i}$$

图8—17表示v'_{ik}的过渡过程，τ_i表示它的时间常数

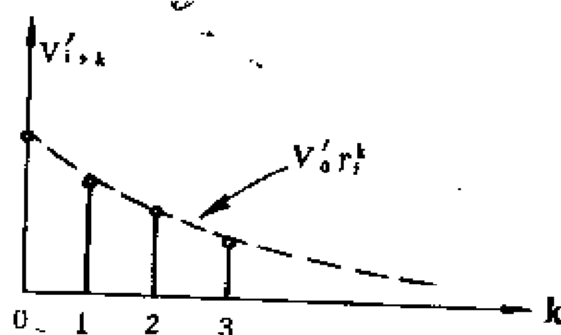


图8-17　适应过程的时间常数

3.最小均方差（LMS）自适应算法及其收敛性　自适应算法的目的是要求得式（8-5-12）的准确或近似的解。直接求解该式，除了要对$n \times n$矩阵求逆外，还要求$n(n+1)/2$个自相关和互相关以便得到R_{xx}和R_{xx}的各元素。当n较大时计算量就很大。在实际问题中输入信号的统计特性可能有缓慢的变化，上述计算还要不断重复进行。事实上，要完全准确地求解式（8-5-12）也是不可能的，因为要准确估计相关矩阵的元素需要无穷的统计样本。

用式（8-5-15）递推计算参数时也要用到R_{xx}和R_{xx}，因此同样存在上述情况。下面提出的求解式（8-5-12）的近似方法是基于最速下降法的LMS算法，它无需计算相关函数，也没有矩阵求逆，甚至不需要平方、平均或微分的运算来求梯度。这种算法的精度受限于样本长度，因为我们只能用有限时间的观察值进行参数估计。

LMS算法的关系式为

$$\varphi_{k+1}=\varphi_k-\mu\hat{\nabla}[E\varepsilon_k^2] \tag{8-5-30}$$

这里估计的梯度$\hat{\nabla}$是根据单个误差平方样本的梯度来计算，即

$$\hat{\nabla}[E\varepsilon_k^2]=\nabla[\varepsilon_k^2]=2\varepsilon_k\nabla[\varepsilon_k] \tag{8-5-31}$$

而由式（8-5-6）有

$$\nabla[\varepsilon_k]=\nabla[x_k-\varphi_k^T X_k]=-X_k$$

将它代入式（8-5-31）得

$$\hat{\nabla}[E\varepsilon_k^2]=-2\varepsilon_k X_k=-2[x_k-\boldsymbol{\varphi}_k^{\mathrm{T}}X_k]X_k \qquad (8\text{-}5\text{-}32)$$

这个梯度估计是无偏的，因为对于某一给定的参数向量$\boldsymbol{\varphi}_k$，梯度估计的期望为

$$\begin{aligned}E[\hat{\nabla}[E\varepsilon_k^2]]&=-2E[X_k(x_k-\boldsymbol{\varphi}_k^{\mathrm{T}}X_k)]\\&=-2[R_{xX}-R_{XX}\boldsymbol{\varphi}_k]=\nabla[E\varepsilon_k^2]\end{aligned}$$

利用梯度估计的式（8-5-32）作参数递推计算时，

$$\boldsymbol{\varphi}_{k+1}=\boldsymbol{\varphi}_k+2\mu\varepsilon_k X_k \qquad (8\text{-}5\text{-}33)$$

这就是 LMS 自适应算法的实际公式，它只需简单的乘法和加法运算，每次递推需 $2n+1$ 次乘法和加法。μ值为一小的常数，其取值范围将在后面讨论。

下面来看 LMS 算法参数向量均值的收敛性，为此假定相邻两次递推间隔时间足够长，以致向量 X_k和X_{k+1}是无关的。由于递推的$\boldsymbol{\varphi}_k$仅仅是X_{k-1}，$X_{k-2},\cdots,X_0$的函数（由式(8-5-33)）。因此，上述假定意味着$\boldsymbol{\varphi}_k$和 X_k无关。对于满足这一条件的平稳过程，可证明当递推次数很大时，$E\boldsymbol{\varphi}_k$ 将收敛于最优解 $\boldsymbol{\varphi}_{LMk}$。

在式（8-5-33）两边取期望可得参数向量期望值的差分方程

$$\begin{aligned}E\boldsymbol{\varphi}_{k+1}&=E\boldsymbol{\varphi}_k+2\mu E[X_k(x_k-X_k^{\mathrm{T}}\boldsymbol{\varphi}_k)]\\&=(I-2\mu R_{XX})E\boldsymbol{\varphi}_k+2\mu R_{xX}\end{aligned} \qquad (8\text{-}5\text{-}34)$$

设$\boldsymbol{\varphi}_0$为初始参数向量，有

$$\begin{aligned}E\boldsymbol{\varphi}_{k+1}&=[I-2\mu R_{XX}]^{k+1}\boldsymbol{\varphi}_0+2\mu\sum_{j=0}^{k}[I-2\mu R_{XX}]^j R_{xX}\\&=[I-2\mu Q^{-1}\Lambda Q]^{k+1}\boldsymbol{\varphi}_0+2\mu\sum_{j=0}^{k}[I-\\&\quad-2\mu Q^{-1}\Lambda Q]^j R_{xX}\\&=Q^{-1}[I-2\mu\Lambda]^{k+1}Q\boldsymbol{\varphi}_0+2\mu Q^{-1}\sum_{j=0}^{k}[I-\\&\quad-2\mu\Lambda]^j QR_{xX}\end{aligned} \qquad (8\text{-}5\text{-}35)$$

现在考察式中的对角阵 $(I-2\mu\Lambda)$，由于它的每个元素的模都小于 1，因此

$$\lim_{k\to\infty}[I-2\mu\Lambda]^{k+1}\to 0$$

即式（8-5-35）的第一项随递推次数增加而消失，而第二项一般收敛到非零的极限，因为根据几何级数有

$$\sum_{i=0}^{\infty}(1-2\mu\lambda_i)^i=\frac{1}{1-(1-2\mu\lambda_i)}=\frac{1}{2\mu\lambda_i}$$

故

$$\lim_{k\to\infty}\sum_{i=0}^{\infty}[I-2\mu\Lambda]^i=\frac{1}{2\mu}\Lambda^{-1}$$

于是式（8-5-35）的极限为

$$\lim_{k\to\infty}E\varphi_{k+1}=Q^{-1}\Lambda^{-1}QR_{xX}=R_{XX}^{-1}R_{xX}=\varphi_{LMS} \quad (8\text{-}5\text{-}36)$$

参数向量对 φ_{LMS} 的收敛性取决于 μ 是否在某一范围内，这个范围就是

$$|1-2\mu\lambda_{max}|<1$$

或

$$0<\mu<1/\lambda_{max} \quad (8\text{-}5\text{-}37)$$

这个条件和梯度按式（8-5-11）计算时（即没有梯度噪声的精确情况）的稳定条件式（8-5-26）是相同的。由于

$$\lambda_{max}\leqslant \mathrm{trace}R_{XX}=E[X_k^TX_k]=\sum_{i=1}^{n}E[x_{k-i}^2] \quad (8\text{-}5\text{-}38)$$

可见收敛的条件也可用

$$0<\mu<1\Big/\sum_{i=1}^{n}E[x_{k-i}^2]$$

实际的自适应递推计算，如果要使结果准确（但过程缓慢）通常

$|\mu| \ll 1 / \sum_{i=1}^{n} E[x_{k-i}^2]$。至于两次递推的输入样本的独立性条件对于收敛性来说并非绝对需要的，但当它们相互有关时会使递推结果的稳态值有较大的误差。

关于LMS算法的过渡过程时间，和前面一样为

$$\tau_i = \frac{1}{2\mu\lambda_i} \qquad i=1,2,\cdots, n$$

如果所有的特征值均相等，则所有参数递推过程的时间常数都一样

$$\tau = 1/2\mu\lambda$$

对于均方误差来说，由于它是参数的二次函数，所以均方误差的过程也是指数形式，而且它的时间常数应为$\tau/2$。

特征值不等时，时间常数的估计要复杂得多，而且实际的参数递推过程也不是光滑的指数形式，因为递推过程总是有噪声（梯度估计噪声），当然，递推进行得越慢，噪声引起的波动也越小。

当过渡过程结束后，LMS算法的残差平方ε_k^2对ε_{min}^2（式（8-5-13））的偏差用相对的平均残差方差超值M表示，它定义为

$$M \equiv \frac{E[\varepsilon_k^2 - \varepsilon_{min}^2]}{\varepsilon_{min}^2}$$

可以证明，当φ_k接近φ_{LMS}时有

$$M = \frac{1}{2} \sum_{i=1}^{n} \frac{1}{\tau_i}$$

当所有时间常数均相等时，M和参数数目成正比，和时间常数成反比，即

$$M = \frac{n}{2\tau}$$

以上讨论都针对平稳过程，LMS 算法也适用于时变的输入过程，然而这时M值除了梯度噪声引起的分量外，还有因参数的真值改变而引起的适应过程的滞后误差。有关时变情况下的讨论可参阅有关研究论文。

4.自适应模型用于预报 在建立自适应 AR 模型时，我们把当时的观察值x_t作为理想值（真值）来与加权和y_t比较。在实时的自适应预报中，用什么作为真值呢？如果要求现在时刻的 l 步预报，显然未来的真值 x_{t+l} 是不知道的，要是知道就无需预报了。为了能够提供一个误差信号来实现计算加权系数的 LMS 算法，我们将 l 步以前的预报结果和现在时刻的值作比较得到的误差来进行自适应建模，然后利用这个模型计算现在对未来的预报。

图8—18是这种算法的原理图，图中"自适应建模"方框的内容和图8—14中的虚线方框的内容相同，只是输入不是x_t而是

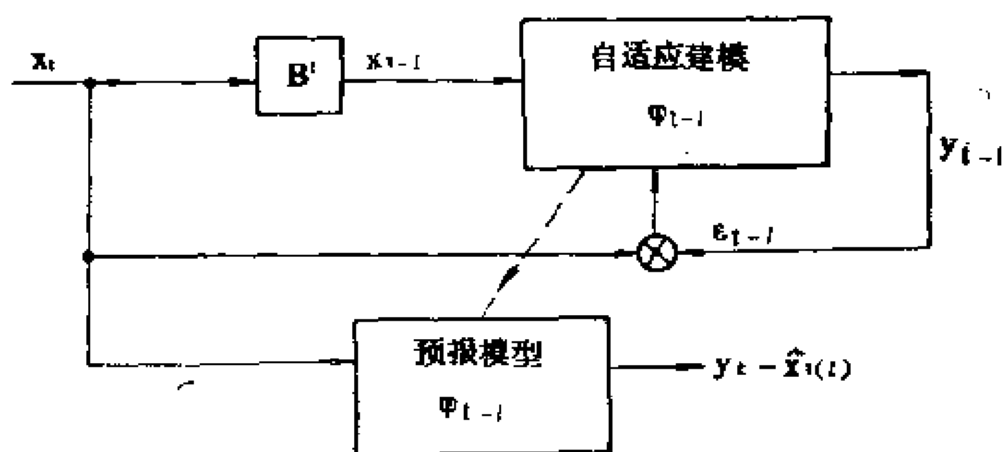


图8-18 自适应模型与预报

x_{t-l}，输出(加权和)记作y_{t-l}，而由x_t和y_{t-l}之差(记作ε_{t-l})作为误差信号控制LMS算法，其递推关系为

$$\begin{aligned}\varphi_{k+1-l}&=\varphi_{k-l}+2\mu\varepsilon_{k-l}X_{k-l}\\&=\varphi_{k-l}+2\mu(x_k-\varphi^T_{k-l}X_{k-l})X_{k-l}\end{aligned}$$

在自适应的过渡过程结束时，由此产生的φ_{t-l}就是能使l步预报误差均方值为最小的AR模型系数，图中的“预报模型”就是把这些系数原样照搬地构成$AR(n)$模型，利用这个模型对x_t，$x_{t-1}, \cdots, x_{t-n}$加权产生y_t作为x_{t+l}的最小方差预报$\hat{x}_t(l)$。即

$$\hat{x}_t(l)=\varphi_{t-l}X_t$$

为了进一步说明时变AR模型参数的自适应递推算法，我们引入一个仿真实例。

调用程序DAGENT产生两个AR(1)序列

$$x_{1t}=0.7x_{1,t-1}+\varepsilon_{1t} \qquad 1\leqslant t\leqslant 200$$

$$x_{2t}=-0.3x_{2,t-1}+\varepsilon_{2t} \qquad 1\leqslant t\leqslant 400$$

由这两个序列合成一个序列$\{x_t\}$：

$$x_t=x_{1t} \qquad 1\leqslant t\leqslant 200$$

$$x_t=x_{2,t-200} \qquad 200<t\leqslant 600$$

调用程序ADAPAR对序列$\{x_t\}$的参数进行自适应递推计算。图8—19中给出了仿真计算结果，(a)、(b)、(c)中分别取μ为0.002，0.005，0.05。从图中我们可以看出参数μ的选取是十分重要的。μ值过小（见(a)）则模型参数的预测值较为平坦，但“跟踪”过于迟缓，不能体现出模型参数的变化。μ值较大（见(c)），则预测值能迅速反映出模型参数的变化，但起伏振荡太大。当取$\mu=0.05$时(见(b))，在$50\leqslant k\leqslant 200$范围内$\varphi_k=0.7$，当$200<k<400$时，模型参数估计处于调整阶段，当$k>400$以后$\varphi_k$趋于模型参数真值$-0.3$。因此取$\mu=0.05$是较为适宜的。

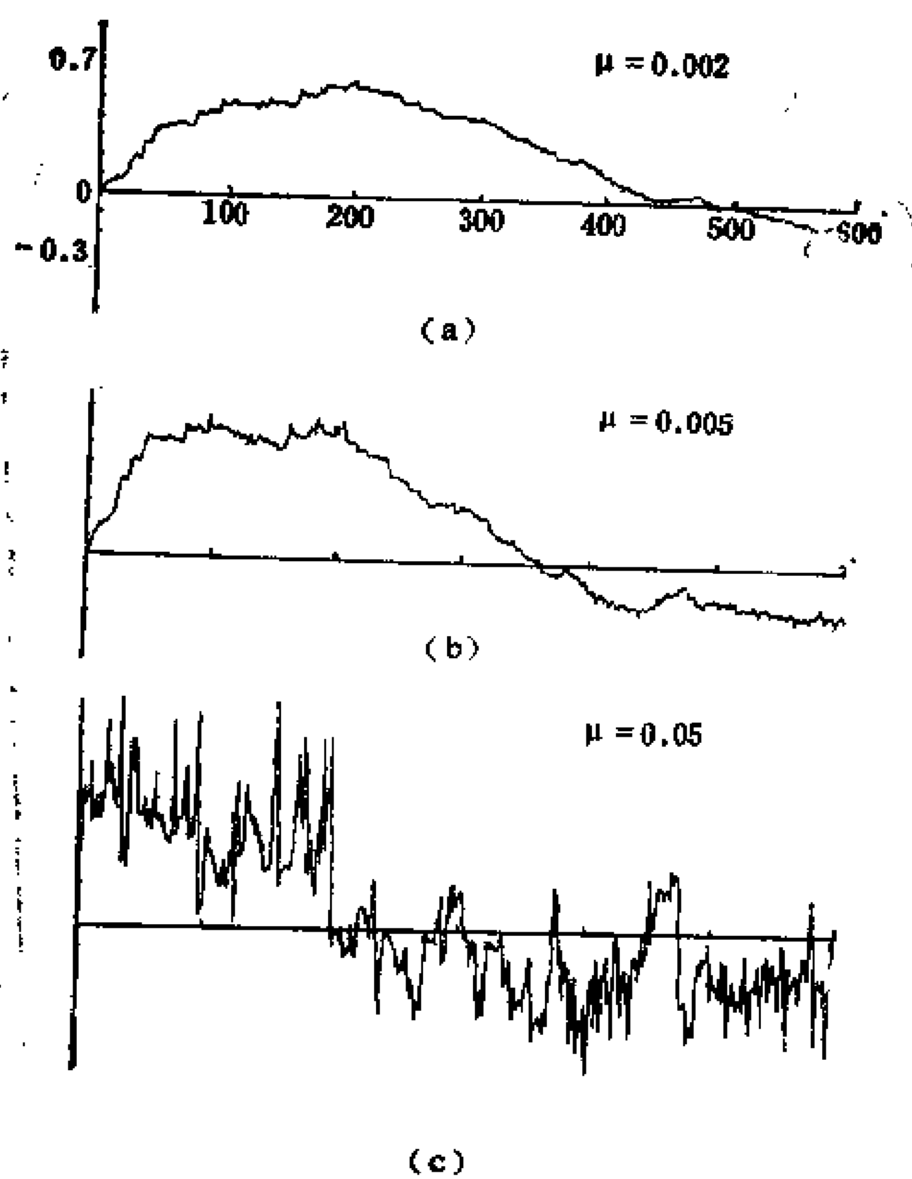


图8-19　自适应AR模型仿真实例

附录一　富氏变换及其算法

频率域函数估计的实际计算工具是建立在富氏变换、特别是离散富氏变换及其快速算法(FFT)的基础上，本附录对它们的性质和基本原理作必要的介绍，以配合第四章和第七章中的有关内容。

1.富氏变换(FT)和离散富氏变换(DFT)

FT和DFT有许多相似之处，为了便于对照，将它们列在一起。关于它们的证明可查阅有关书籍。

FT的定义——复单值函数x(t)的富氏变换为

$$\mathscr{F}[x(t)] = X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} dt \qquad (\text{Ⅰ}-1)$$

其中t和f为实变量，若t为时间(秒)，则f为频率(Hz)，$j=\sqrt{-1}$。

X(f)的富氏反变换为

$$\mathscr{F}^{-1}\left[X(f)\right] = x(t) = \int_{\infty}^{-\infty} X(f) e^{j2\pi ft} df \qquad (\text{Ⅰ}-2)$$

除在x(t)的间断点处积分给出$[x(t^-)+x(t^+)]/2$外，$X(f)$的$\mathscr{F}^{-1}$变换便是x(t)。

DFT的定义——N个复数构成的序列$\{x_0, x_1, \cdots, x_{N-1}\}$的离散富氏变换是另一个由N个复数构成的序列$\left\{X_0, X_1, \cdots, X_{N-1}\right\}$，其中

$$X_k=\sum_{n=0}^{N-1}x_n{}^{-j2\pi nk/N}\qquad k=0,\ 1,\ \cdots,\ N-1\qquad (\text{I}-3)$$

而序列$\left\{X_0,\ X_1,\ \cdots,\ X_{N-1}\right\}$的离散富氏反变换(*IDFT*)则给出原序列$\{x_0,\ x_1,\ \cdots,\ x_{N-1}\}$其算式为

$$x_n=\frac{1}{N}\sum_{k=0}^{N-1}X_k e^{j2\pi nk/N}\qquad n=0,\ 1,\ \cdots N-1\qquad (\text{I}-4)$$

引入简便的符号

$$W_N=e^{-j2\pi/N}\qquad (\text{I}-5)$$

则式(Ⅰ-3)和(Ⅰ-4)可写成

$$X_k=\sum_{n=0}^{N-1}x_n W_N^{nk}\qquad k=0,1,\cdots,N-1\qquad (\text{I}-6)$$

$$x_n=\frac{1}{N}\sum_{k=0}^{N-1}X_k W_N^{-nk}\qquad n=0,1,\cdots,N-1\qquad (\text{I}-7)$$

式(Ⅰ-7)可证明如下：利用

$$\sum_{k=0}^{N-1}W_N^{qk}=\begin{cases}N & q=0,\pm N,\pm 2N,\cdots\\ 0 & q\neq 0,\pm N,\pm 2N,\cdots\end{cases}\qquad (\text{I}-8)$$

将式(Ⅰ-6)代入式(Ⅰ-7)右边得

$$\frac{1}{N}\sum_{k=0}^{N-1}\ \sum_{p=0}^{N-1}x_p W_N^{pk}W_N^{-nk}=\frac{1}{N}\sum_{p=0}^{N-1}x_p\ \sum_{k=0}^{N-1}W_N{}^{(p-n)k}=x_n$$

根据*DFT*的定义，由于W_N^n以N为周期，即$W_N^m=W_N^{m\bmod N}$，其中m为任意整数

$$m\bmod N=\begin{cases}m-N\lfloor m/N\rfloor & N=\pm 1,\pm 2,\cdots\\ m & N=0\end{cases}\qquad (\text{I}-9)$$

$\lfloor\cdot\rfloor$表示取小于或等于符号内的值的最大整数。可见

$$X_k=X_{k\bmod N}\qquad k=0,\pm 1,\pm 2,\cdots\qquad (\text{I}-10)$$

$$x_n = x_{n \bmod N} \qquad n=0, \pm 1, \pm 2, \cdots$$

因此在任何方程式中，如果x或X的下标为负或大于N-1，都应取为它的*mod*N的值。例如当N=4时，$x_4=x_0$，$x_5=x_1$，…。

对于序列$\{\cdots, x_{-2}, x_{-1}, x_0, x_1, x_2, \cdots\}$，当$x_{-n}=x_n$时通常称为偶序列，$x_{-n}=-x_n$时称为奇序列，$x_{-n}=x_n^*$时称为埃尔米特(Hermitian)序列$(n=0, \pm 1, \pm 2, \cdots)$。目前情况下，由于$(N-n) \bmod N=-n$故$x_{N-n}$，等于$x_{-n}$，所以若序列$\{x_0, x_1, \cdots, x_{N-1}\}$中$x_n=x_{N-n}$则称为偶序列，$-x_n=x_{N-n}$时称为奇序列，而$x_n^*=x_{N-n}$时称为埃尔米特序列。$(n=1, 2, \cdots \lfloor N/2 \rfloor)$

下面列出有关*FT*和*DFT*的主要性质，为简便将“富氏变换对”用“↔”记号表示，如$x(t) \leftrightarrow X(f)$，$x_n \leftrightarrow X_k$分别表示x(t)和X(f)之间的FT关系和$x_n$和$X_k$之间的*DFT*关系。

（1） 线性

FT：$c_1x(t)+c_2y(t) \leftrightarrow c_1X(f)+c_2Y(f)$ （Ⅰ-11）

DFT：$c_1x_n+c_2y_n \leftrightarrow c_1X_k+c_2Y_k$ （Ⅰ-12）

其中c_1, c_2为任意复数

（2）绕原点的翻转

FT：$x(-t) \leftrightarrow X(-f)$ （Ⅰ-13）

DFT：$x_{-n} \leftrightarrow X_{-k}$ （Ⅰ-14）

（3）复共轭

FT：$x^*(t) \leftrightarrow X^*(-f)$ （Ⅰ-15）

$x^*(-t) \leftrightarrow X^*(f)$ （Ⅰ-16）

DFT：$x_n^* \leftrightarrow X^*_{-k}$ （Ⅰ-17）

$x_{-n}^* \leftrightarrow X_k^*$ （Ⅰ-18）

（4） 对称性

偶函数、奇函数或实的偶函数经正的或反的富氏变换后都不会改变原来的奇偶性，但实的奇函数变换后是虚的奇函数。表达式如下：

$$FT:\ x(t)=x(-t)\leftrightarrow X(f)=X(-f) \tag{Ⅰ-19}$$

$$x(t)=-x(-t)\leftrightarrow X(f)=-X(-f) \tag{Ⅰ-20}$$

$$x(t)=x^*(t)=x(-t)=x^*(-t)\leftrightarrow X(f)=X^*(-f)$$
$$=X(-f)=X^*(f) \tag{Ⅰ-21}$$

$$x(t)=x^*(t)=-x(-t)=-x^*(-t)\leftrightarrow X(f)=X^*(-f)$$
$$=-X(-f)=-X^*(f) \tag{Ⅰ-22}$$

$$DFT:\ x_n=x_{-n}\leftrightarrow X_k=X_{-k} \tag{Ⅰ-23}$$

$$x_n=-x_{-n}\leftrightarrow X_k=-X_{-k} \tag{Ⅰ-24}$$

$$x_n=x_n^*=x_{-n}=x_{-n}^*\leftrightarrow X_k=X_{-k}^*=X_{-k}=X_k^* \tag{Ⅰ-25}$$

$$x_n=x_n^*=-x_{-n}=-x_{-n}^*\leftrightarrow X_k=X_{-k}^*=-X_{-k}=-X_k^* \tag{Ⅰ-26}$$

由于任意函数$x(t)$可表为偶函数$e(t)=[x(t)+x(-t)]/2$和奇函数$o(t)=[x(t)-x(-t)]/2$，不难证明实函数的FT（或实序列的DFT）必具有偶实部和奇虚部（这种函数称为埃尔米特函数）。即

若$x(t)$为实，则$X(f)$为埃尔米特函数$(X(-f)=X^*(f))$

若x_n为实，则X_k为埃尔米特函数$(X_{-k}=X_k^*)$

若$X(f)$为实，则$x(t)$为埃尔米特函数$(x(-t)=x^*(t))$

若X_k为实，则x_n为埃尔米特函数$(x_{-n}=x_n^*)$

（5） 位移

$$FT:\ x(t-b)\leftrightarrow e^{-j2\pi fb}X(f) \tag{Ⅰ-27}$$

$$e^{j2\pi tb}x(t)\leftrightarrow X(f-b) \tag{Ⅰ-28}$$

b为任意实数

$$DFT:\ x_{n-p}\leftrightarrow W^{pk}X_k \tag{Ⅰ-29}$$

$$W^{-pn}x_n\leftrightarrow X_{k-p} \tag{Ⅰ-30}$$

（6） 面积与求和

$$FT:\ \int_{-\infty}^{\infty}x(t)dt=X(0) \tag{Ⅰ-31}$$

$$\int_{-\infty}^{\infty} X(f)df = x(0) \qquad (\text{I}-32)$$

$$DFT: \sum_{n=0}^{N-1} x_n = X_0 \qquad (\text{I}-33)$$

$$\sum_{k=0}^{N-1} X_k = N x_0 \qquad (\text{I}-34)$$

（7） 卷积

$$FT: x(t)*y(t) \leftrightarrow X(f)Y(f) \qquad (\text{I}-35)$$

$$x(t)y(t) \leftrightarrow X(f)*Y(f) \qquad (\text{I}-36)$$

这里*表示线性卷积，即

$$x(t)*y(t) = \int_{-\infty}^{\infty} x(u)y(t-u)du = \int_{-\infty}^{\infty} x(t-u)y(u)du \qquad (\text{I}-37)$$

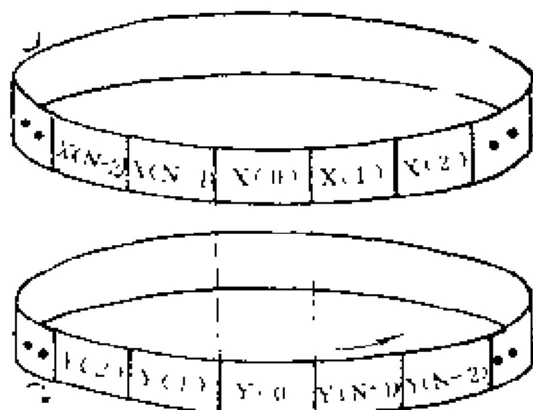
$$DFT: \sum_{p=0}^{N-1} x_p y_{n-p} \leftrightarrow X_k Y_k \qquad (\text{I}-38)$$

$$N x_n y_n \leftrightarrow \sum_{p=0}^{N-1} X_p Y_{k-p} \qquad (\text{I}-39)$$

这里离散卷积又称循环卷积，两个有限序列(x_n和y_n)长度是相同的。如果用两个标有各自数值的圆环对在一起。循环卷积相当于把一个序列的N个值按逆时针顺序标在环上（见图附Ⅰ—1），另一序列的N个值按顺时针方向写在环上，二环对着的各个值分别相乘，然后将这N个乘积相加得出某一循环卷积的值，每次把第二个环反时针（或第一个环顺时针）转过一格，按同样方法相乘求和可得到另一个卷积值。

（8）相关

$$FT: x(t)*y^*(-t) \leftrightarrow X(f)Y^*(f) \qquad (\text{I}-40)$$



图附Ⅰ-1　循环卷积

$$x(t)y^*(t)\leftrightarrow X(f)*Y^*(-f) \tag{Ⅰ-41}$$

这里相关定义为

$$\int_{-\infty}^{\infty}x(u)y^*(u-t)du=\int_{-\infty}^{\infty}x(u+t)y^*(u)du=x(t)*y^*(-t)$$

(Ⅰ-42)

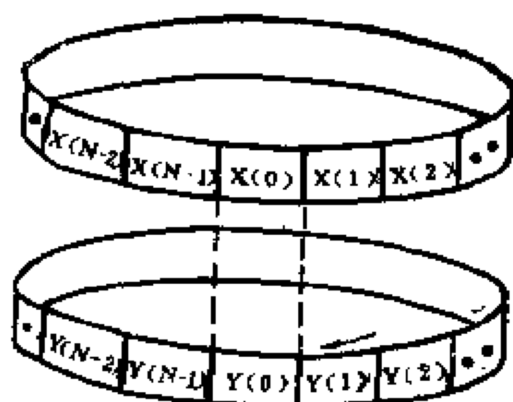
t为相关的延迟

$$DFT:\ \sum_{p=0}^{N-1}x_{p+n}y_p^*\leftrightarrow X_kY_k^* \tag{Ⅰ-43}$$

$$Nx_ny_n^*\leftrightarrow\sum_{p=0}^{N-1}X_{p+k}Y_p^* \tag{Ⅰ-44}$$

这种相关也称循环相关，它也要求二序列的长度相同，其形象解释和循环卷积类似，只是第二个序列y_n也按反时针方向写在环上，各相关值是由y_n序列顺时针（或x_n序列反时针）转动一格而得到的不同乘积产生（见图附Ⅰ-2）。

如果要用循环相关（或卷积）的运算求出两个长度不等（如一个是N_1，另一个是N_2）的序列的线性相关（或卷积），则两个序列的长度都要改为(N_1+N_2-1)，为此要在两个序列末尾各



图附 I-2　循环相关

添(N_2-1)和(N_1-1)个零。

(9)能量

$$FT: \int_{-\infty}^{\infty} x(t)y^*(t)dt = \int_{-\infty}^{\infty} X(f)Y^*(f)df \qquad (\text{I}-45)$$

当$x(t)=y(t)$时，$|X(f)|^2$($x(t)$的能量谱)的面积等于$|x(t)|^2$的面积($x(t)$的总能量)，这就是帕色瓦尔(Parseval)定理。

$$DFT: \sum_{n=0}^{N-1} x_n y_n^* = \frac{1}{N}\sum_{k=0}^{N-1} X_k Y_k^* \qquad (\text{I}-46)$$

当$x_n=y_n$时，这就是离散序列的帕色瓦尔定理。

2.采 样 定 理

在实际情况下，时间函数和它在频率域中的变换都是连续函数，而在计算处理中又往往要加以离散化，因此采样是不可避免的，如何经采样而又不丢失应有的信息是采样定理所要指出的主要内容。

（1） 时间域采样定理——若x(t)为单值、频带宽度有限($f_s Hz$)的时间函数，则当采样间隔取为$\triangle=1/(2f_s)$时，可按下式精确地复现$x(t)$

$$x(t)=\sum_{n=-\infty}^{\infty}x(n\Delta)\frac{\sin[2\pi f_s(t-n\Delta)]}{2\pi f_s(t-n\Delta)} \qquad (\text{Ⅱ}-1)$$

证明如下：

因x(t)带宽为f_s，故有

$$x(t)=\int_{-f_s}^{f_s}X(f)e^{j2\pi ft}df$$

令$t=n\Delta$，则上式可写作

$$x(n\Delta)=\int_{-f_s}^{f_s}X(f)e^{j\pi nf/f_s}df$$

将$X(f)$在$-f_s<f<f_s$内展成复富氏级数，即

$$X(f)=\sum_n A_n e^{-j\pi nf/f_s} \qquad |f|<f_s$$

其中 $A_n=\frac{1}{2f_s}\int_{-f_s}^{f_s}X(f)e^{j\pi nf/f_s}df=\frac{x(n\Delta)}{2f_s}$

用$\widetilde{X}(f)$表示按$X(f)$重复的周期函数，它是

$$\widetilde{X}(f)=\sum_n\frac{x(n\Delta)}{2f_s}e^{-j\pi nf/f_s}$$

将$\widetilde{X}(f)$在$|f|<f_s$范围的部分截取出来，再作富氏反变换就应当是x(t)。这种频域截取等于$\widetilde{X}(f)$和矩形函数$Q(f)$相乘，$Q(f)$定义为

$$Q(f)=\begin{cases}1 & |f|<f_c\\ 1/2 & f=f_c\\ 0 & |f|>f_c\end{cases}$$

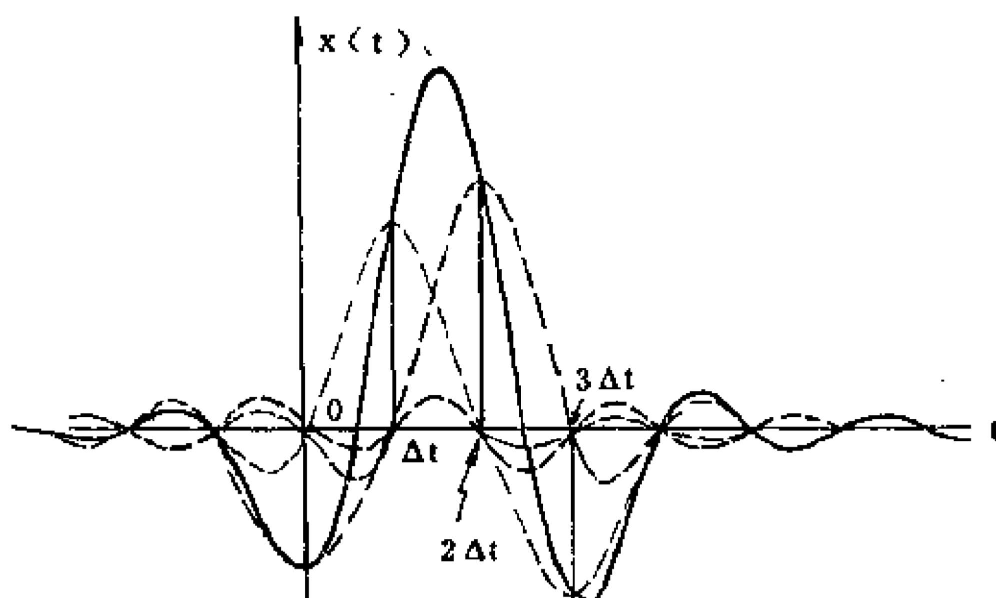
$$x(t)=\mathscr{F}^{-1}\left[\widetilde{X}(f)Q(f)\right]=x(n\Delta)*\mathscr{F}^{-1}[Q(f)],$$ 由于

$$\mathscr{F}^{-1}[Q(f)]=\frac{\sin(2\pi f_c t)}{2\pi f_c t}$$

故得式(Ⅱ-1)。

时域采样定理说明，虽然通过一组给定的采样点可以有无穷多个函数，但只有符合式(Ⅱ-1)的才是频带为f_c的$x(t)$，采样频率至少是$2f_c$。如果关于f_c的限制不满足，那末用式(Ⅱ-1)给出的函数是一种近似。

图附Ⅰ-3表示某时间函数的采样[…，0，0，$x(0)$，$x(\Delta)$，$x(2\Delta)$，$x(3\Delta)$，0，0，…]与按式(Ⅱ-1)得出的$x(t)$曲线。



图附Ⅰ-3　时间域采样

(2)　频率域采样定理——若$x(t)$总长度为T秒（当$|t|\geq T/2$时$x(t)=0$），其富氏变换$X(f)$可以由频率间隔为$\Delta_f=1/T$的

X采样值来唯一地确定，即

$$X(f)=\sum_{k=-\infty}^{\infty}X\left(\frac{k}{T}\right)\frac{\sin[\pi T(f-k/T)]}{\pi T(f-k/T)} \qquad (\text{Ⅱ}-2)$$

该定理的证明和时间域的采样定理类似，从略。从频域采样定理可以看出，当$T=N\Delta$时Δ_f应不大于$1/(N\Delta)$。若$\Delta=1$，则在$-\frac{1}{2}$至$\frac{1}{2}$的频率范围内要取N个以上的采样才能获得必要的频域信息。

3. FT，DFT，FS(富氏级数)间的关系

设$x(t)$是$0<t<T$内的复单值函数，在其他t处$x(t)=0$，其中$T=N\Delta$(Δ为采样间隔，N为正整数)。以$X(f)$表$x(t)$的FT，由于$X(f)$的反变换在$x(t)$的间断点处给出$[x(t^-)+x(t^+)]/2$，考虑到采样点处可能有间断，故令$t=n\Delta$处$x(t)$的采样值定义为

$$x(n\Delta)=[x(n\Delta^-)+x(n\Delta^+)]/2 \qquad n=\cdots,-2,-1,0,1,2\cdots$$

考虑到　$t\leqslant 0$和$t\geqslant T$时$x(t)=0$故

$$x(0)=[x(0^-)+x(0^+)]/2=x(0^+)/2$$

$$x(N\Delta)=[x(N\Delta^-)+x(N\Delta^+)]/2=x(N\Delta^-)/2$$

今按下式构成N项序列$\{x_0,x_1,\cdots,x_{N-1}\}$

$$x_n=\begin{cases}x(0)+x(N\Delta)=[x(0^+)+x(N\Delta^-)]/2 & n=0\\ x(n\Delta)=[x(n\Delta^-)+x(n\Delta^+)]/2 & n=1,2,\cdots N-1\end{cases} \qquad (\text{Ⅱ}-1)$$

这个序列用$x(t)$和单位脉冲的无穷序列　$p(t)=\sum_{n=-\infty}^{\infty}\delta(t-n\Delta)$的乘积表示，即（见图附Ⅰ-4）

$$x(t)p(t)=\sum_{n=0}^{N}x(n\Delta)\delta(t-n\Delta)$$

由于

$$\mathscr{F}[p(t)]=P(t)=\frac{1}{\Delta}\sum_{n=-\infty}^{\infty}\delta\left(f-\frac{n}{\Delta}\right) \tag{Ⅱ-2}$$

根据卷积关系式(Ⅰ-36)，有

$$\mathscr{F}[x(t)p(t)]=X(f)*P(f)=\frac{1}{\Delta}\sum_{i=-\infty}^{\infty}X\left(f-\frac{i}{\Delta}\right) \tag{Ⅱ-3}$$

然而，同样可以证明$\mathscr{F}[x(t)p(t)]$正是序列$\{x_n\}$的DFT，因为

$$\mathscr{F}[x(t)p(t)]=\int_{-\infty}^{\infty}\left[\sum_{n=0}^{N}x(n\Delta)\delta(t-n\Delta)\right]e^{-j2\pi ft}dt$$

$$=\sum_{n=0}^{N}\left[x(n\Delta)\int_{-\infty}^{\infty}\delta(t-n\Delta)e^{-j2\pi ft}dt\right]=\sum_{n=0}^{N}x(n\Delta)e^{-j2\pi fn\Delta} \tag{Ⅱ-4}$$

显然$F[x(t)p(t)]$是周期为$1/\Delta$的周期函数，令$f=k/(N\Delta)$，根据定义式(Ⅱ-1)，式(Ⅱ-4)最后一个等式可改写成

$$\sum_{n=0}^{N-1}x_n e^{-j2\pi kn/N}$$

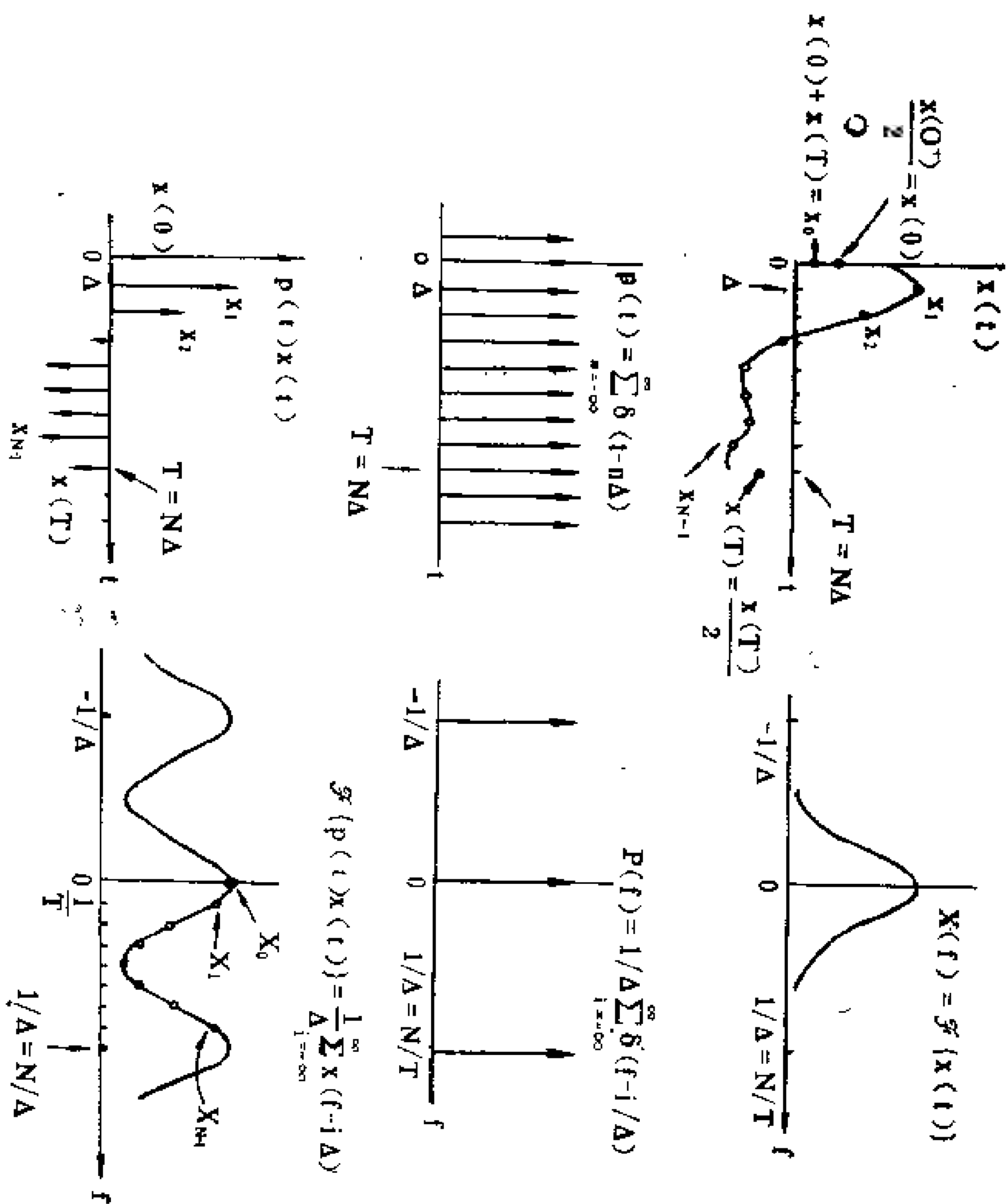
这和式(Ⅰ-3)的定义一致，因此

$$X_k=\frac{1}{\Delta}\left[\sum_{i=-\infty}^{\infty}X(f-\frac{i}{\Delta})\right]_{f=k/(N\Delta)}\quad k=0,1,\cdots N-1 \tag{Ⅱ-5}$$

该式说明，X_k是由无穷多个相隔$1/\Delta$的$X(f)$叠加而成。如果N个Δ正好充满整个$x(t)$长度，即$T=N\Delta$，则式(Ⅱ-5)可以写成

$$X_k=\frac{1}{\Delta}X\left(\frac{k}{T}\right)+\frac{1}{\Delta}\sum_{\substack{i=-\infty\\ i\neq 0}}^{\infty}X\left(\frac{k-iN}{T}\right)\quad k=0,\pm 1,\cdots,\pm\lfloor N/2\rfloor \tag{Ⅱ-6}$$

注意，$k=\lfloor N/2\rfloor+1,\cdots,N-1$时的$X_k$实际上代表负频率时



图附 I-4　FT和DFT关系的说明

的X_k值，式(Ⅱ-6)右边第二项和式代表$|i|>1$的所有$X\left(\frac{k}{T}-\frac{i}{\Delta}\right)$项都要叠加到第一项来构成$X_k$，这是“混叠”现象的数学表达，当$X(f)$带宽限于$|f|<1/(2\Delta)$时上述第二项为零，没有频率混叠。

下面来看FT和FS的关系。由于FS是对周期函数的展开，因此要先对$x(t)$作周期延拓，即假定$\tilde{x}(t)$在$0<t<T$范围内和$x(t)$相同，而在此范围以外按周期T重复，利用式(2-4-3)(2-4-4)有

$$\tilde{x}(t)=\sum_{k=-\infty}^{\infty}A_k e^{j2\pi kt/T} \qquad (Ⅱ-7)$$

$$A_k=\frac{1}{T}\int_0^T \tilde{x}(t)e^{-j2\pi kt/T}dt \qquad k=0,\pm1,\pm2,\cdots \qquad (Ⅱ-8)$$

由此可得周期函数的FT和复富氏级数系数A_k的关系为（见图附Ⅰ-5）

$$\mathscr{F}\left[\tilde{x}(t)\right]=\int_{-\infty}^{\infty}\sum_{k=-\infty}^{\infty}A_k e^{j2\pi kt/T}e^{-j2\pi ft}dt$$

$$=\sum_{k=-\infty}^{\infty}A_k\int_{-\infty}^{\infty}e^{-j2\pi\left(f-\frac{k}{T}\right)t}dt$$

$$=\sum_{k=-\infty}^{\infty}A_k\delta\left(f-\frac{k}{T}\right) \qquad (Ⅱ-9)$$

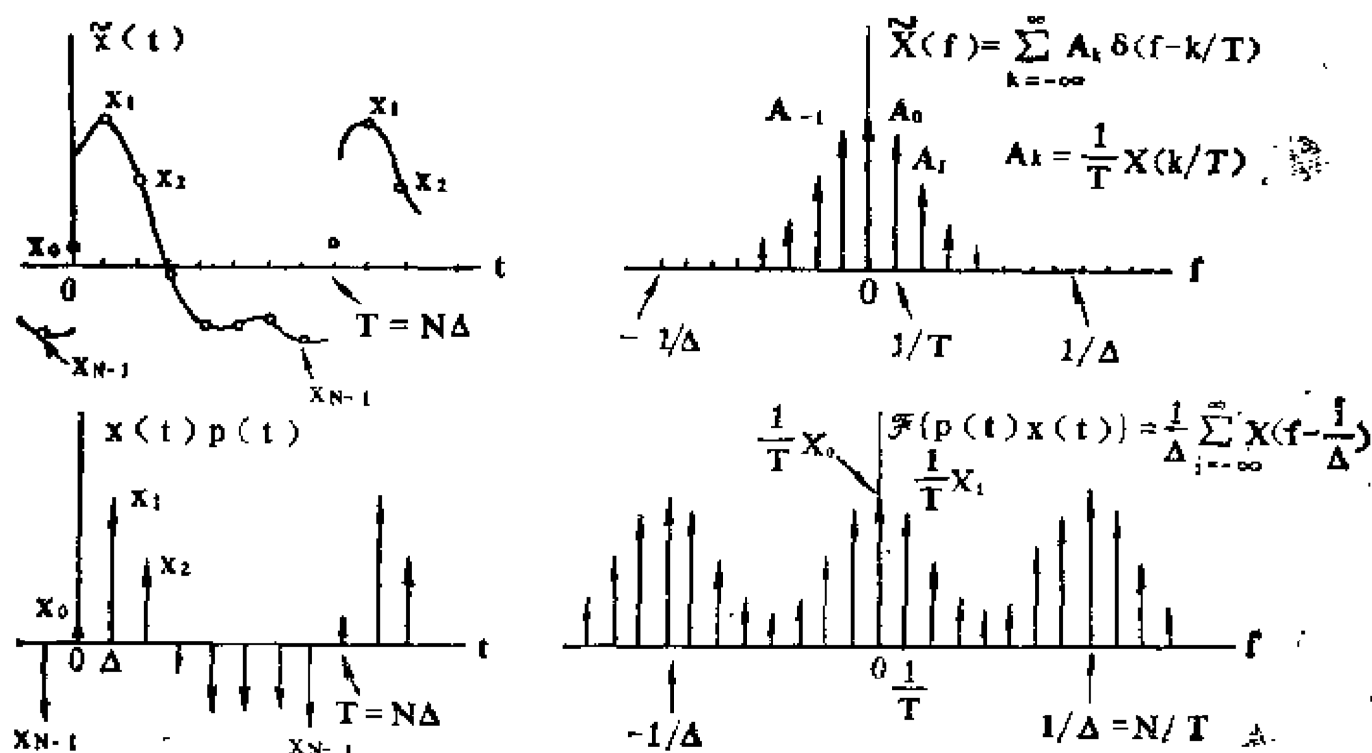
对照式(Ⅱ-8)和式(Ⅰ-1)有

$$A_k=\frac{1}{T}X\left(\frac{k}{T}\right) \qquad k=0,\pm1,\pm2,\cdots \qquad (Ⅱ-10)$$

将它代入式(Ⅱ-6)可得DFT和FS系数的关系

$$X_k = NA_k + N\sum_{\substack{i=-\infty \\ i\neq 0}}^{\infty} A_{k+iN} \quad k=0,\pm 1,\cdots,\pm\lfloor N/2\rfloor \quad (Ⅱ-11)$$

显然，如果$\tilde{x}(t)$可展成有限项复富氏级数，则当$|k|\geq N/2$时$A_k=0$，这时式(Ⅱ-11)没有第二项和式。



图附Ⅰ.3 DFT和FS关系的说明

4. DFT的快速算法—FFT

按式(Ⅰ-3)直接计算N点复序列的DFT约需N^2次运算（一次运算是指一次复数乘法和加法）。1965年库利(*Cooley*)和杜奇(*Tukey*)发表论文介绍了当N是2的幂时如何以大约$N\log_2 N$次的运算求得DFT，这个成果是很轰动的，因为当N很大时，$N\log_2 N$比N^2要小得多。例如当$N=2^{10}=1024$时$N^2/(N\log_2 N)=$

102.4。这使谱函数的计算大为省时，此外这种算法在计算机内存的占用上也很节省，这使得频域方法分析在许多情况下变得十分优越、人们把只需$N\log_2 N$次运算的DFT算法称为快速富氏变换(FFT)。

FFT具体算法很多，它们在编程和硬件实现上各有特点，这里只就$N=2^m$时常用的两种方法说明算法原理。

（1） 时序重排法

一次分解：将式(Ⅰ-6)写成偶数i和奇数i的和式：

$$X_k=\sum_{n=0}^{N-1} x_n W_N^{nk}=\sum_{n=0}^{(N/2)-1} x_{2n} W_N^{2nk}+\sum_{n=0}^{(N/2)-1} x_{2n+1} W_N^{(2n+1)k} \tag{Ⅳ-1}$$

由于$W_{N/2}=W_N^2$

$$X_k=\sum_{n=0}^{(N/2)-1} x_{2n} W_{N/2}^{nk}+W_N^k\sum_{n=0}^{(N/2)-1} X_{2n+1} W_{N/2}^{nk} \tag{Ⅳ-2}$$

注意到上式右边第一项是$N/2$点由x_n的偶数点的值构成（记作$a_n=x_{2n}$，$n=0,1,\cdots,N/2-1$)的DFT，而第二项是$N/2$点由x_n的奇数点的值构成（记作$b_n=x_{2n+1}$，$n=0,1\cdots,N/2-1$）的DFT，分别用A_k和B_k表示这两个DFT，则

$$X_k=A_k+W_N^k B_k \qquad k=0,1,\cdots N-1 \tag{Ⅳ-3}$$

由于DFT的周期性，$A_{k+N/2}=A_k$，$B_{k+N/2}=B_k$，所以N点DFT——X_k可由两个$N/2$点DFT——A_k和B_k求得。如果再考虑到$W_N^{k+N/2}=-W_N^k$，则式(Ⅳ-3)等价于一对方程式：

$$X_k=A_k+W_N^k B_k \qquad k=0,1,\cdots,N/2-1 \tag{Ⅳ-4}$$

$$X_{k+N/2}=A_k-W_N^k B_k \qquad k=0,1,\cdots,N/2-1$$

可见当$N/2$个A_k和B_k已知时，按式(Ⅳ-4)求N点X_k只要N次复数加法和$N/2$次复数乘法，因为每个k的乘积$W_N^k B_k$对上述二式可以共用。这里把$W_N^k=1$和j这些特殊情况也计入乘法次数。

二次分解：通过$N/4$点的DFT计算N/2点DFT——A_k和B_k。令$c_n \equiv a_{2n} = x_{4n}$和$d_n \equiv a_{2n+1} = x_{4n+2}(n=0,1,\cdots,\frac{N}{4}-1)$，以$C_k$，$D_k$分别表示$N/4$点($c_n$和$d_n$)的DFT，则

$$A_k = C_k + W_N^{2k} D_k \qquad k=0,1,\cdots N/4-1$$

$$A_{k+N/4} = C_k - W_N^{2k} D_k \quad k=0,1,\cdots,N/4-1 \qquad (\text{IV}-5)$$

类似地，以E_k和F_k表示$e_n \equiv b_{2n} = x_{4n+1}$和$f_n \equiv b_{2n+1} = x_{4n+3}$ ($n=0,1,\cdots,N/4-1$)的$N/4$点DFT，则有

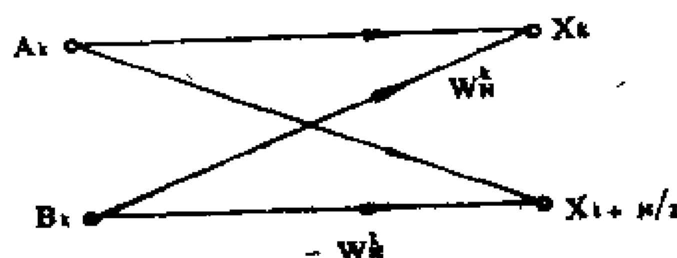
$$B_k = E_k + W_N^{2k} F_k \qquad k=0,1,\cdots,N/4-1$$

$$B_{k+N/4} = E_k - W_N^{2k} F_k \qquad k=0,1,\cdots N/4-1 \quad (\text{IV}-6)$$

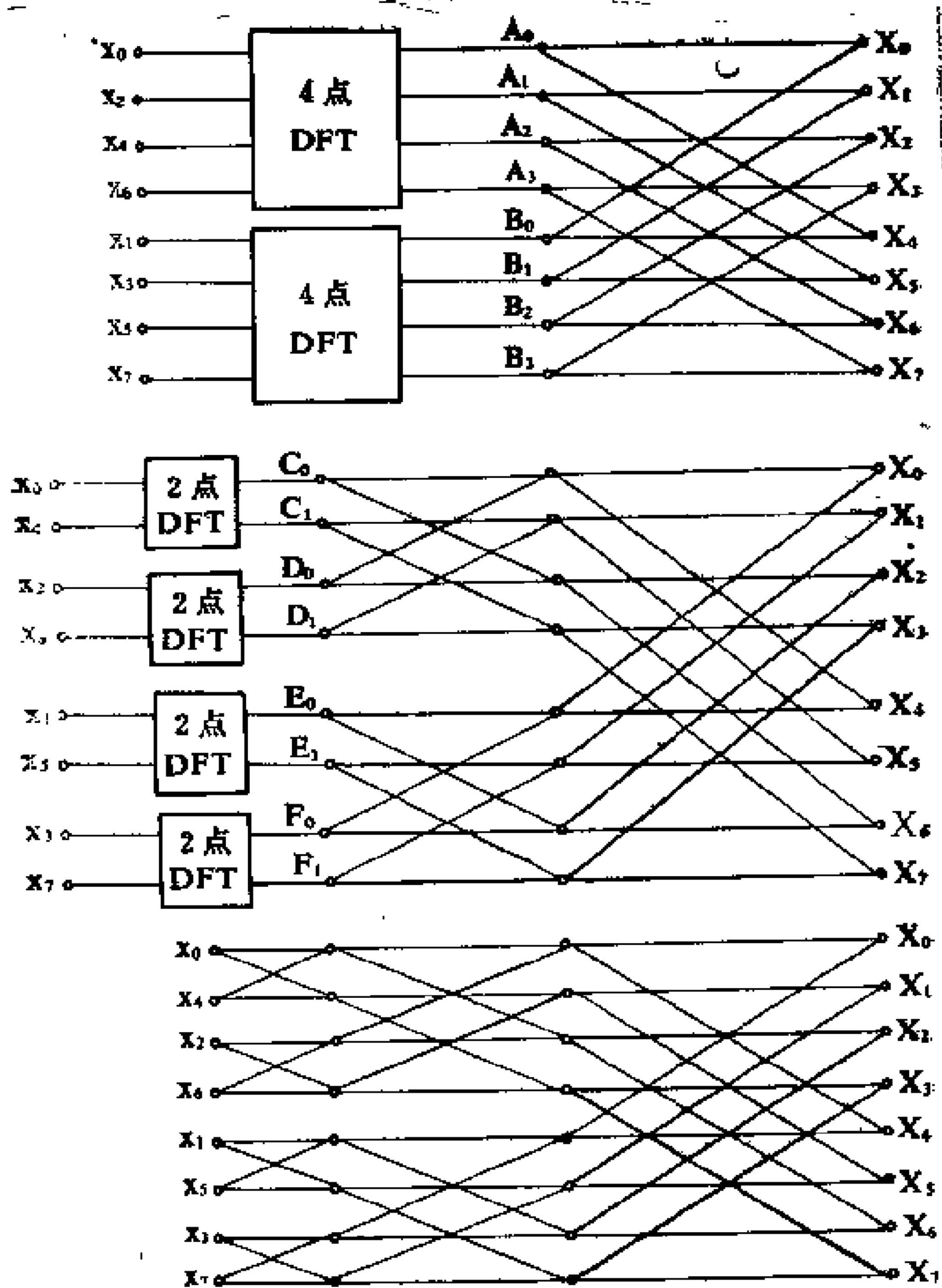
若C_k，D_k已知，按式(Ⅳ-5)计算全部A_k需要$N/2$次复数加法和$N/4$次复数乘法。计算B_k的运算次数也一样。因此计算A_k和B_k总共需要N次复数加法和$N/2$次复数乘法，总次数和第一次相同。

由于$N=2^m$，当分解到第m次($m=\log_2 N$)时对应的是1点DFT，也就是x_n本身，而全部DFT计算共需$N\log_2 N$次复数相加和$(N/2)\log_2 N$次复数相乘。

实际计算DFT是从数据序列x_n开始，所以求取X_k的步骤和上述分解的叙述顺序相反。直观上，FFT计算的基本结构呈"蝴蝶型"如图附Ⅰ-6所示。图附Ⅰ-7给出$N=8=2^3$时的三步分解所组成的运算结构。由图可见，X_k在右边各节点自上而下按自然顺序排列，而序列x_n在左边各节点上出现的顺序要按二进制的逆序重新排列，其方法是先将n写成二进制数，然后将该数的各位顺序反过来再换算成十进制数，表附Ⅰ-1是n由0到7的举例说明。



图附Ⅰ-6　2点DFT的蝶形图



图附1-7 时序重排法的演化过程说明

表附Ⅰ-1 二进制位的逆序

n	二进制表示	逆序的二进制表示	逆序的二进制数所对应的十进制数
0	0 0 0	0 0 0	0
1	0 0 1	1 0 0	4
2	0 1 0	0 1 0	2
3	0 1 1	1 1 0	6
4	1 0 0	0 0 1	1
5	1 0 1	1 0 1	5
6	1 1 0	0 1 1	3
7	1 1 1	1 1 1	7

在按时序重排法编制FFT计算程序时，先把x_n的自然顺序按其二进制位的逆序重排，然后依蝶形结构进行计算.

（2） 频序重排法

如果把DFT公式（Ⅰ-6)分成前后两个序列，即

$$X_k=\sum_{n=0}^{N/2-1} x_n W_N^{nk}+\sum_{n=0}^{N/2-1} x_{n+N/2} W_N^{(n+N/2)k}$$

$$=\sum_{n=0}^{N/2-1}[x_n+(-1)^k x_{n+N/2}]W_N^{nk} \qquad k=0,1,\cdots N-1$$

(Ⅳ-7)

把X_k分成偶数点和奇数点，即

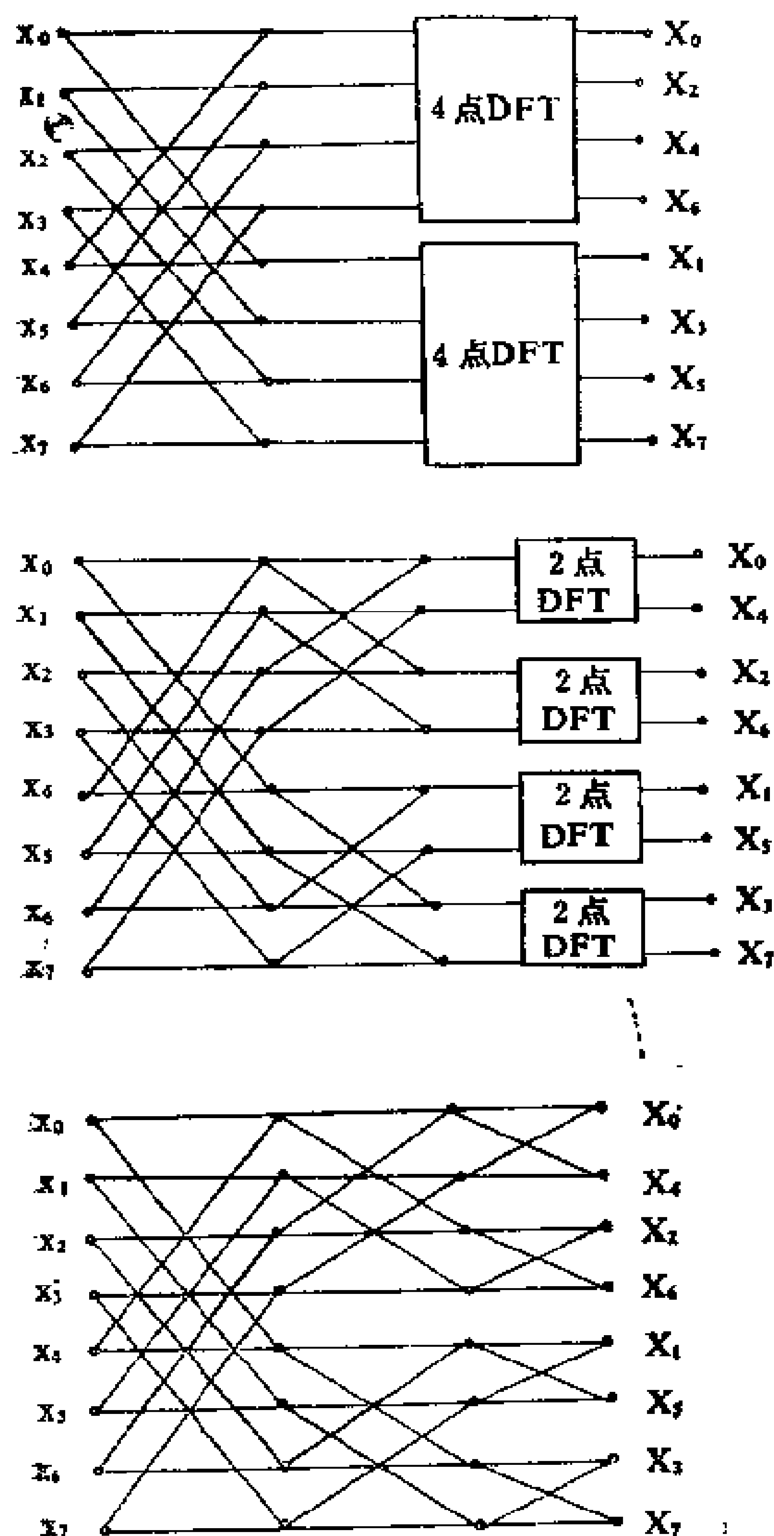
$$X_{2k}=\sum_{n=0}^{N/2-1}[x_n+x_{n+N/2}]W_{N/2}^{nk} \qquad k=0,1,\cdots N/2-1$$

$$X_{2k+1}=\sum_{n=0}^{N/2-1}[(x_n-x_{n+N/2})W_N^n]W_{N/2}^{nk}$$

$$k=0,1,\cdots,N/2-1$$

(Ⅳ-8)

可见 X_{2k} 是 N/2点序列 $a_n = x_n + x_{n+N/2}$ $(n=0,1,\cdots,N/2-1)$ 的DFT，而 X_{2k+1} 是 N/2 点序列 $b_n = (x_n - x_{n+N/2})\, W_N^n$



图附 I-8 频序重排法的演化过程说明

$(n=0,1,\cdots,N/2-1)$的DFT。由x_n和$x_{n+N/2}$求a_n和b_n的运算结构也是蝴蝶型，需要N次复数加法和N/2次复数乘法。

进一步可以将N/2点DFT分解，其做法和上述相同，一直分解到1点DFT即x_n本身为止，共需$m=\log_2 N$步分解，也要做$N\log_2 N$次复数加法和$(N/2)\log_2 N$次复数乘法。图附Ⅰ-8为$N=8$时的算法结构图。可以看出，数据序列x_n的自然顺序不变，但计算结果的X_k按二进位逆序排列，我们可以按表附Ⅰ-1所示方法恢复它的自然顺序。

虽然上面讨论的是求离散富氏变换式(Ⅰ-3)〔或式(Ⅰ-6)〕的快速算法，但实际上FFT也可用来求离散富氏反变换，如果将式(Ⅰ-4)取共轭

$$x_n^*=\frac{1}{N}\sum_{k=0}^{N-1}X_k^*e^{-j2\pi nk/N}=\sum_{k=0}^{N-1}(X_k^*/N)e^{-j2\pi nk/N}$$

$$=\mathrm{DFT}(X_k^*/N) \qquad (\text{Ⅳ}-9)$$

可见，当已给X_k时可以通过正变换和取共轭求得它的反变换x_n，因此算法和程序对两种变换是可共用的。

5. 两个实数序列的同时变换

在谱估计中，特别是采用直接法时，需要对大量实序列求富氏变换。由于FFT可以处理复数序列，如果把两个长度相同的实序列分别作为复序列的实部和虚部，则可利用FFT算法以较短的时间完成全部实序列的变换。

设x_n和$y_n(n=0,1,\cdots,N-1)$是两个实序列，以其为实部和虚部构成一个复序列z_n：

$$z_n=x_n+jy_n \qquad (\text{Ⅴ}-1)$$

根据线性性质，其富氏变换为

$$Z_k=X_k+jY_k \qquad (\text{Ⅴ}-2)$$

而由复共轭关系式（Ⅰ-17）知

$$z_n^* = x_n - jy_n \leftrightarrow Z_{-k}^* = X_k - jY_k \qquad (\text{V}-3)$$

由式（Ⅴ-2）和（Ⅴ-3）并考虑到$Z_{-k}^* = Z_{N-k}^*$，可得

$$\left.\begin{aligned} X_k &= \frac{1}{2}\left[Z_k + Z_{N-k}^*\right] \\ Y_k &= \frac{1}{2j}\left[Z_k - Z_{N-k}^*\right] \end{aligned}\right. \qquad k = \begin{cases} 0, 1, \cdots, \dfrac{N}{2} & N\text{偶} \\ 0, 1, \cdots, \dfrac{N-1}{2} & N\text{奇} \end{cases}$$

$$(\text{V}-4)$$

注意，X_k和Y_k并非就是Z_k的实部和虚部，因为X_k和Y_k本身一般是复数，而且对于实数x_n和y_n而言，它们的DFT具有偶的实部和奇的虚部。

设Z_k的实部为R_k，虚部为I_k，即$Z_k = R_k + jI_k$，则一般说，R_k和I_k（实函数）既非偶函数亦非奇函数，但可将其表为偶函数和奇函数的叠加。

$$R_k = \underbrace{\left[\frac{R_k}{2} + \frac{R_{-k}}{2}\right]}_{\text{偶}} + \underbrace{\left[\frac{R_k}{2} - \frac{R_{-k}}{2}\right]}_{\text{奇}} \equiv R_k^e + R_k^o \qquad (\text{V}-5)$$

$$I_k = \underbrace{\left[\frac{I_k}{2} + \frac{I_{-k}}{2}\right]}_{\text{偶}} + \underbrace{\left[\frac{I_k}{2} - \frac{I_{-k}}{2}\right]}_{\text{奇}} \equiv I_k^e + I_k^o$$

这里R_k^e，I_k^e和R_k^o，I_k^o分别表示式中括号内的偶函数和奇函数。因此有

$$Z_k = R_k^e + R_k^o + jI_k^e + jI_k = [R_k^e + jI_k^o] + j[I_k^e - jR_k^o] \qquad (\text{V}-6)$$

根据式（Ⅴ-2）中X_k，Y_k本身实部和虚部的奇偶性质，同式（Ⅴ-6）对照可见

$$X_k = R_k^e + jI_k^o = \tfrac{1}{2}[(R_k + R_{-k}) + j(I_k - I_{-k})] \qquad (\text{V}-7)$$

$$Y_k = I_k^e - jR_k^o = \tfrac{1}{2}[(I_k + I_{-k}) - j(R_k - R_{-k})]$$

式(Ⅴ-7)关系可以在不同场合加以利用。如果要求两个实序列的DFT，则可同时变换得出Z_k，然后按式(Ⅴ-7)进行分离求得各自的DFT—X_k和Y_k。另一种情况是处理一个长的实序列，可以先将它分成两个等长的短序列，同时求出两个序列的DFT之后再将结果组合起来得到整个序列的DFT。

在直接法的谱估计中，对分段的加窗数据求各自的周期图时，可以利用上述方法同时变换两段数据，例如以相邻两段实数据分别作为复数z_i的实部和虚部，并以X_k和Y_k分别表示各自的DFT。可以证明

$$
\begin{aligned}
|X_k|^2+|Y_k|^2 &= \frac{1}{2}\left(Z_k Z_k^* + Z_{-k} Z_{-k}^*\right) \\
&= \frac{1}{2}\left(Z_k Z_k^* + Z_{M-k} Z_{M-k}^*\right) \qquad (\text{Ⅴ}-8)
\end{aligned}
$$

(M为每段的数据个数)，可见由复序列z_i的DFT可以算出两段实序列周期图之和。对于L段数据只需经过$L/2$次类似的运算，其结果的总和将可得出分段周期图的叠加。它的总体平均即为所需的谱估计。

6. 相关函数的快速计算

FFT也可用来计算样本相关函数

$$
\hat{R}(r)=\frac{1}{N}\sum_{n=0}^{N-|r|-1} x_n x_{n+r} \quad 0\leqslant r\leqslant M-1,\ M\leqslant N \quad (\text{Ⅵ}-1)
$$

固然可以按照循环相关的变换关系式(Ⅰ-43)，在增添必要的零值数据后，通过$X_k X_k^*$求得x_n的线性自相关，但当N很大时需要作长序列的DFT计算，其效率不高甚至是不现实的。本节方法是利用M(r的变化范围)通常比N小得多的特点，将长序列重新组合以加速运算。

设$N=LM$，并考虑到在$0\leqslant n\leqslant N-1$区间外$x_n=0$，将式(Ⅵ-1)上限改为$N-1$，改写成分段相关之和

$$\hat{R}(r)=\frac{1}{N}\left[\sum_{n=0}^{M-1}x_n x_{n+r}+\sum_{n=M}^{2M-1}x_n x_{n+r}+\cdots+\sum_{n=(L-1)M}^{LM-1}x_n x_{n+r}\right]$$

$$=\frac{1}{N}\sum_{i=1}^{L}\left[x_{(i-1)M}x_{(i-1)M+r}+x_{(i-1)M+1}x_{(i-1)M+1+r}+\cdots\right.$$

$$\left.+x_{(i-r)M+M-1}x_{(i-1)M+M-1+r}\right)$$

$$=\frac{1}{N}\sum_{i=1}^{L}\sum_{n=0}^{M-1}x_{n+(i-1)M}x_{n+(i-1)M+r} \qquad (\text{VI}-2)$$

定义 $v_i(r)=\sum_{n=0}^{M-1}x_{n+(i-1)M}\,x_{n+(i-1)M+r} \qquad 0\leqslant r\leqslant M-1$

$$(\text{VI}-3)$$

则 $\hat{R}(r)=\frac{1}{N}\sum_{i=1}^{L}v_i(r) \qquad 0\leqslant r\leqslant M-1 \qquad (\text{VI}-4)$

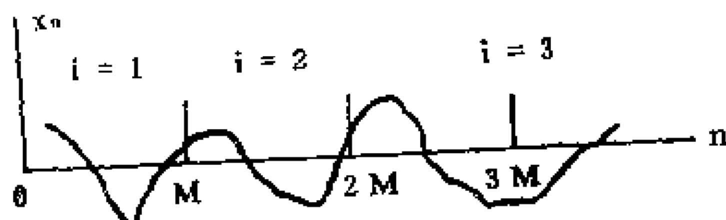
不难看出，$v_i(r)$表示把x_n分成L段($i=1,2,\cdots,L$)后各段的x_n和x_{n+r}乘积的总和。例如（见图附Ⅰ-9）

$$v_1(r)=\sum_{n=0}^{M-1}x_n x_{n+r}$$

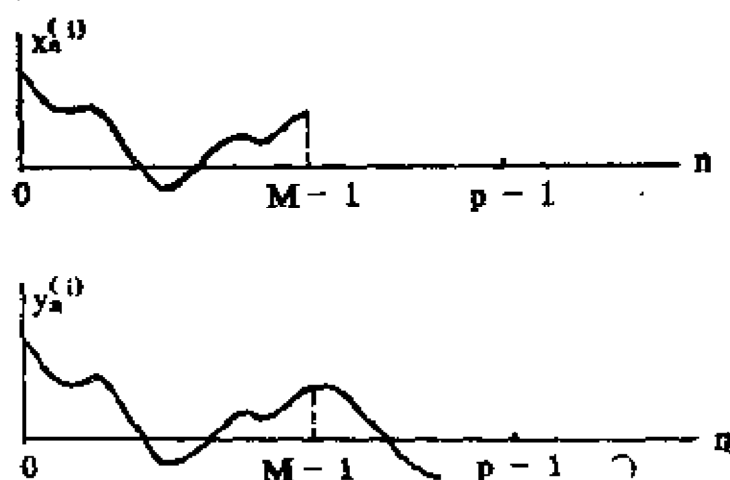
$$v_2(r)=\sum_{n=0}^{M-1}x_{n+M}x_{n+M+r}$$

$$\cdots$$

每个v_i的计算可以利用循环相关。为此，对第i段定义（见图附Ⅰ-10）



图附Ⅰ-9　数据分段



图附Ⅰ-10　求相关的两段数据

$$x_n^{(i)}=\begin{cases}x_{n+(i-1)M} & 0\leqslant n\leqslant M-1\\ 0 & M\leqslant n\leqslant P-1\end{cases}\qquad (\text{Ⅵ}-5)$$

$$y_n^{(i)}=x_{n+(i-1)M}\qquad 0\leqslant n\leqslant P-1$$

这里要求$P\geqslant 2M-1$。由于$x_n^{(i)}$和$y_n^{(i)}$的循环相关在$0\leqslant r\leqslant M-1$的范围内等于线性相关$v_i(r)$，因此先求式(Ⅵ-5)的$P$点DFT——$X_k^{(i)}$和$Y_k^{(i)}$，计算$V_k^{(i)}=X_k^{(i)}Y_k^{(i)*}$，则$\sum\limits_{i=1}^{L}V_k^{(i)}$ $(k=0,1,\cdots P-1)$的P点离散反变换就是$\sum\limits_{i=1}^{L}v_i(r)$。再除以$N$便得$\hat{R}(r)$。

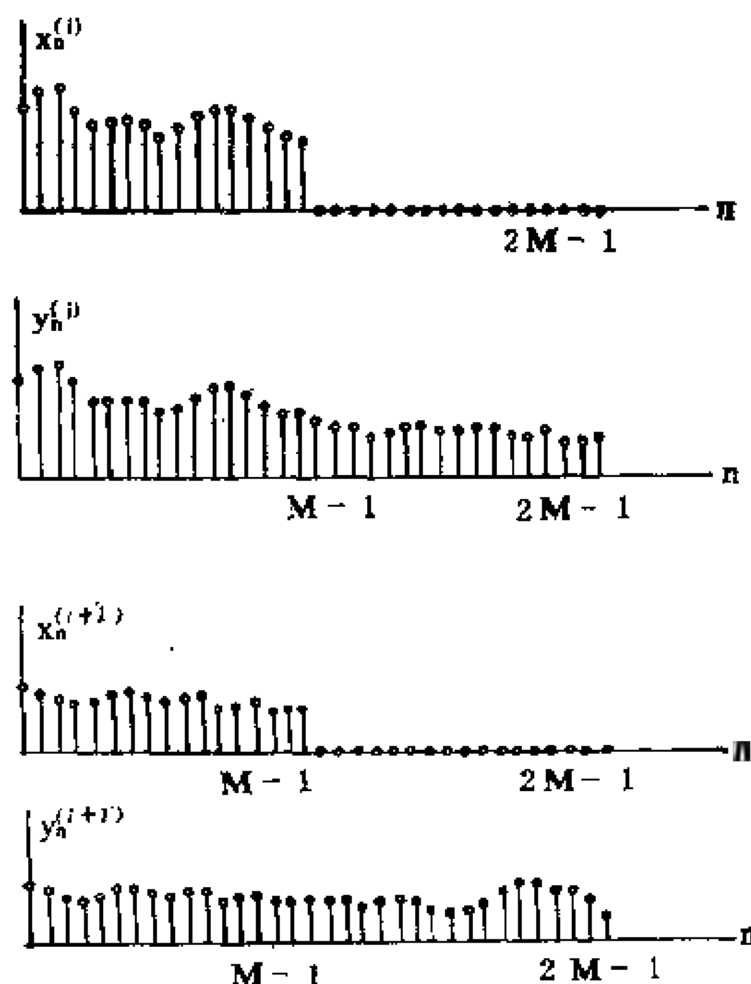
若取$P=2M$，可进一步简化$\hat{R}(r)$的计算，图附Ⅰ-11表示这时的$x_n^{(i)}$，$y_n^{(i)}$和$x_n^{(i+1)}$，$y_n^{(i+1)}$，显然

$$y_n^{(i)}=x_n^{(i)}+x_{n-M}^{(i+1)}$$

故有（由式(Ⅰ-29)并注意到$W_P^M=W_{2M}^M=-1$）

$$Y_k^{(i)}=X_k^{(i)}+(-1)^kX_k^{(i+1)}\qquad k=0,1,\cdots,2M-1$$

可见$Y_k^{(i)}$无需另外的FFT计算，而且$X_k^{(i)}$和$X_k^{(i+1)}$的计算还可



图附Ⅰ-11　当取P＝2M时$x_n^{(i+1)}$的前半段等于$y_n^{(i)}$的后半段

利用两段序列同时变换的方法。总之，$\hat{R}(r)$的快速计算方法步骤为：

（1）构造序列

$$x_n^{(1)}=\begin{cases}x_n & 0\leqslant n\leqslant M-1\\ 0 & M\leqslant n\leqslant 2M-1\end{cases}$$

计算$P=2M$点DFT—$X_k^{(1)}$，令$A_k^{(0)}=0$

（2）构造序列

$$x_n^{(i+1)}=\begin{cases}x_{n+iM} & 0\leqslant n\leqslant M-1\\ 0 & M\leqslant n\leqslant 2M-1\end{cases}\qquad i=1,2,\cdots L$$

计算

$$A_k^{(i)}=A_k^{(i-1)}+X_k^{(i)}[X_k^{(i)*}+(-1)^k X_k^{(i+1)*}]$$

$$0\leqslant k\leqslant 2M-1,\quad i=1,2,\cdots,L$$

(3)　令$V_k=A_k^{(L)}$，并求其$2M$点DFT反变换$v(r)$，则

$$\hat{R}(r)=\frac{1}{N}v(r)\qquad 0\leqslant r\leqslant M$$

以上总共要求L次变换($X_k^{(i)}$)和一次反变换。

附 录 二

时间序列分析与建模程序说明

本附录所列全部程序均可在IBM-PC/XT微机或可与其兼容的微机上执行，操作系统为PC-DOS版本2。

程序运行所需的原始数据（如时间序列的观察值等）预先存在用户给定的数据文件中，通过执行OPEN和READ语句按自由格式读入，其他必要的输入参数通过人机对话由键盘打入，运行结果均显示在屏幕上，同时也可打印。用户可按需要改变或附加其他的输出方式。

1. 随机过程数据产生程序（DAGENT）

1.1 原理简介

过程的数学模型是由两个确定性的正弦信号和一个随机过程叠加而成，即

$$X_t = A_1\cos(2\pi f_1 t) + A_2\cos(2\pi f_2 t) + x_t$$

其中振幅A_1，A_2和频率f_1，f_2是用户给定的。x_t可以是白噪声过程，也可以是任何ARMA过程，即

$$x_t - \phi_1 x_{t-1} - \cdots - \phi_n x_{t-n} = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_m \varepsilon_{t-m}$$

当x_t为白噪声时，其概率密度函数可在指数分布（即$\lambda e^{-\lambda x}$，$x>0$；$\lambda>0$，其他x处为零）、均匀分布（即$1/(b-a)$，$a\leq x\leq b$，其他x处为零）和正态分布三者中任选。当x_t取为ARMA过程时则按正态分布。

为产生上述过程，程序中利用乘同余法产生具有均匀分布、指数分布或正态分布的随机数。

1.2 主程序功能和参数说明

根据给定的正弦振幅、频率以及随机过程x_t的形式，产生所需的 X_t 过程。

当x_t为白噪声时，有三种概率分布任选——指数分布（这时要给定参数 λ）；均匀分布（这时要给定界限 a 和 b）；正态分布（这时要给定均值和方差）。

当x_t为ARMA过程时，需要给定正态分布ε_t过程的均值和方差以及自回归和滑动平均的阶数和系数。

本程序产生的离散X_t值最多为2048个。

〔输入量〕

A1，A2——正弦振幅

F1，F2——正弦频率（Hz）

IA——随机过程选择

IA＝1时x_t为白噪声过程

IA＝2时x_t为ARMA过程

NQB——需要产生的X_t序列数据总数

ID——白噪声分布选择

ID＝1 指数分布

ID＝2 均匀分布

ID＝3 正态分布

A——分布的参数，ID＝1时表指数分布的λ；

ID＝2时表均匀分布的下限；ID＝3时表正态分布的方差。

B——分布的参数，ID＝2时表均匀分布的上限；ID＝3时表正态分布的均值。

AVG——x_t（白噪声）或ε_t的均值。

VAR——x_t（白噪声）或ε_t的方差。

〔输出量〕

X——产生的时序数据。

1.3 子程序说明

子程序RANDOM用乘同余法产生在（a,b）区间上均匀分布、指数分布或正态分布的随机数。调用本子程序一次就产生一个随机数。

子程序 RANDOM中参数A，B，ID 的意义同主程序说明，其他输入参数有

G——计数器

RNI——先放初值，最后结果作为下一次调用时的初值。

SN——存放各种分布的随机数。

子程序SML用于产生ARMA过程

NAR——自回归阶数n

NMA——滑动平均阶数m

PHI——自回归系数ϕ_1，ϕ_2，…，ϕ_n

THETA——滑动平均系数θ_1，θ_2，…，θ_m

AT——ARMA过程的冲量ε_t

X——ARMA过程x_t的数据

NOB——x_t数据总数。

1.4 程序运行及举例

产生 ARMA（2,2）过程：$x_t+1.45x_{t-1}+0.6x_{t-2}=\varepsilon_t-0.2\varepsilon_{t-1}-0.1\varepsilon_{t-2}$，其中$\varepsilon_t\sim N(0,1)$，$x_t$点数为500。

```
C>DAGENT OUTPUT
AMPLITUDES OF TWO SINEWAVES---A1,A2=?
0. 0.
FREQUENCIES OF TWO SINEWAVES---f1, f2(IN Hz)=?
0. 0.
TYPE OF ADDED PROCESS---1=WHITE NOISE, 2=ARMA
PROCESS
2
AVERAGE OF WHITE NOISE---AVG=?
```

```
0.
VARIANCE OF WHITE NOISE---VAR=?
1.
NO OF SIMULATE DATA---N=?
500
AR ORDER---n=?      MA ORDER---m=?
 2 2
AR PARAMETERS---PHI(1), · · · , PHI(n)=?
-1.45 -.6
MA PARAMETERS---THETA(1), · · · , THETA(m)=?
.2 .1
XAVG=    -.00951    XVAR=    10.52717
StoP - Program terminated.
```

产生的数据存在OUTPUT文件中

C>TYPE OUTPUT

```
 -.1701    .6223  -2.0506   4.3436  -4.1803   1.3236    .1115  -1.2540
 2.1768  -3.9047   4.0851  -3.0279    .2498   2.7849  -4.9234   7.6107
-8.1568   6.1458  -4.4211   3.5178  -2.6406   1.7964   -.5257    .7500
 -.2202   -.1913    .3651  -1.4043   2.6195  -2.4961   1.5320   -.0905
 -.7753   2.4584  -3.2544   2.8821   -.4190  -1.9767   3.9213  -3.7172
 2.9764  -2.7698   1.0530   1.8178  -3.0775   2.3383  -3.4126   5.6488
-5.3757   4.3419  -3.5373   2.1517   -.7994    .5193  -1.0724   1.0609
  .7859   -.0257  -2.4821   3.9014  -4.0342    .9321    .1642    .8113
 -.3931  -1.7455   3.8803  -4.7149   6.0356  -7.1367   5.7075  -3.8086
 1.4526   1.5476  -2.9750   3.8689  -3.1656   2.5557  -3.1518   4.1026
-5.6610   5.7964  -4.5550   2.1905    .1061    .2217    .0423   -.6762
  .7991   -.3734   -.1225   3.2779  -5.9753   8.1316  -9.1606   9.6901
-9.7383   7.6811  -5.2235   4.1631  -2.6022    .0614   2.8997  -4.7651
 2.9876    .1874  -1.0202    .8642  -1.7002   -.2415   1.3108  -1.3403
 2.3202  -3.3729   4.8333  -3.3285   1.9862   -.4460  -1.6922   3.0877
-3.6466   3.1584  -1.1625    .5248   2.2985  -5.2136   5.6367  -3.5829
 2.0069   -.6773   -.9105    .7553  -1.3334   1.6633  -1.6635   1.8002
-2.1591   2.1863  -3.0751   1.5179   3.1077  -4.8771   2.9112  -1.7634
  .8978    .6501  -3.4177   4.7817  -4.2658   4.6300  -3.3317    .5857
  .0677  -1.4408   3.1071  -4.1558   3.2446  -1.5647  -1.3280   4.1678
-4.9091   4.2303  -3.0524   1.0009  -1.8292   2.7882  -2.3733    .4407
```

.9473	−2.9451	3.1233	−3.0621	2.1308	−1.3199	.2955	.1283
.1009	.2016	−1.1356	.9689	−.5274	2.8042	−4.4156	4.9261
−6.4549	6.8508	−6.7378	5.2133	−3.2525	1.6403	−1.3157	1.5743
−3.0609	3.0992	1.0202	−3.3766	.6396	2.6670	−4.6081	5.4867
−4.0078	1.5777	.8261	−1.7471	1.8042	−1.8213	1.0435	−1.9173
4.6616	−7.1173	9.0040	−7.8854	5.1314	−2.0052	.0644	−.2862
.3∞62	1.8089	−3.8893	3.6843	−2.0857	.5111	−.0730	−.0057
1.3192	−1.3477	−.0129	.4526	−.5818	1.1906	−.6320	−.6822
2.0637	−1.6125	1.5134	−1.3303	2.0239	−1.8646	−.3090	.2210
−.8677	2.1753	−2.1700	1.7959	−2.2241	3.1199	−3.6777	3.6985
−2.6628	2.1202	−1.1140	1.1177	−1.6968	2.0491	−3.8453	4.2645
−4.0311	3.5786	−4.4346	3.7856	−2.3830	2.4907	−3.0719	3.1234
−2.0758	2.2399	−2.7293	2.9975	−3.3146	2.4999	−2.3780	.8931
.5725	−.1057	−1.6108	−.0451	.9374	−.8774	1.2132	−1.9090
3.6678	−4.7034	5.1422	−3.8520	3.5949	−5.2243	4.6482	−4.0161
3.1284	−3.2497	3.9072	−1.8707	.3338	2.8315	−4.1209	3.6090
−2.9956	2.3996	−1.9367	.5574	1.5867	−2.3210	1.7400	−1.3340
2.3588	−4.1321	4.6106	−3.4828	2.0428	−.9987	−.2574	1.1328
−1.2392	−.3706	−1.0126	2.4018	−2.8828	4.4892	−5.7385	4.5714
−.8745	−1.0393	2.9131	−4.0144	4.5346	−3.4534	1.9839	−1.4752
−.8190	1.1823	−.5397	−.4249	1.0249	−3.6572	3.8442	−2.9729
1.8121	−.9503	1.2926	−1.1439	1.2850	−2.0368	2.8447	−2.4744

1.8727	—1.3149	.1972	—.8797	.9393	—1.9642	.5420	—.0622
—1.2379	3.2792	—4.8284	4.4343	—2.9396	2.7685	—2.3506	2.2688
—2.7024	1.8728	1.0806	—3.0465	2.4551	.2151	—2.7798	4.3995
—3.4631	2.3941	—1.4653	—.8190	3.4182	—1.8308	4.8456	—5.9657
5.2755	—1.7364	.5116	—1.2359	2.2730	—1.5855	—.4926	.9781
—.5000	—1.6943	2.1843	—.6856	—.5160	1.1321	—1.5430	—.5096
1.5561	—1.0167	.3829	1.2921	—1.5042	—.0425	.7168	.1883
—.6986	.4329	—.8051	2.1899	—1.7976	—.0006	1.2940	—.7825
—1.9850	4.4187	—4.6873	4.0574	—4.9349	4.4710	—3.9291	2.7308
—1.0198	—.2003	1.5177	—2.2202	1.6567	—.7144	.2577	1.0269
—1.2632	2.2382	—4.3216	5.8523	—7.4511	8.1647	—8.4136	8.7334
—8.6225	6.8613	—4.6247	2.8947	—1.0216	—2.1177	8.8533	—6.2452
6.7218	—5.9193	3.5841	1.2205	—3.8226	4.6863	—6.8227	7.4118
—6.7885	5.8034	—3.7423	—.6450	2.7430	—3.3158	3.0849	—4.3898
5.2941	—4.6923	3.5498	—2.4659	1.6546	—1.1605	.1098	1.9746
2.4486	3.0066	—2.9473	1.4402	—.6316	—.2867	1.0609	—.5429
.0164	—1.3006	1.9258	—.7114	.0213	.2056	—1.5912	2.7261
—1.3336	—.4226	.8108	—.2625	—.8449	1.3179	—1.1686	1.5276
.7896	—1.1971	3.5665	—6.3815	8.4881	—8.6118	7.0007	—3.8464
1.1473	.4677	—1.4551	2.2697				

2. 样本均值和样本方差的计算（SMSV）

2.1 原理简介

见第一章1.2节

2.2 主程序功能和参数说明

对观察序列$\{x_t\}$，用多级算法计算其样本均值，用移去均值的直接算法计算其样本方差。

本程序容许最多样本个数为2048。

〔输入量〕

X——观察序列$\{x_t\}$。

N——观察序列的数据个数。

〔输出量〕

EX——样本均值。

SIGM——样本方差。

2.3 程序运行及举例

建立数据文件SMSV.O：

1.,3.,-2.,-1.,0.,1.,2.,-3.,

程序运行结果

```
C>SMSV SMSV.0
  NUMBER OF DATA N= ?
8
AVERAGE=      .1250        VARIANCE=      3.609
Stop - Program terminated.
```

3. 直方图的计算和正态性检验（PDFH）

3.1 原理简介

见第一章1.2和1.3两节

3.2 主程序功能和参数说明

根据给定的直方图计算区间下界、子区间宽度和数目(NX)，对输入数据绘制直方图，并以该数据序列的样本方差和均值计算

相应的正态分布且打印在同一图上。

为了进行正态性的χ^2拟合优度检验，程序给出样本χ^2分布，用户根据它和自由度为NX，显著水平为α的$\chi^2_{NX,\alpha}$值的对比，判断数据分布的正态性。

〔输入量〕

X——分析的数据

NDATA——数据总个数

XBEG——直方图计算区间的下界

DELX——直方图子区间宽度

NX——子区间数目（包括小于计算区间下界和大于计算区间上界的两个子区间）

MR——是否重新计算的标志（*MR*=1 重新计算）

〔输出量〕

HIST——落入各子区间内的数据个数。

CURVE——相应正态分布的直方图数据。

XEND——直方图计算区间的上界，它等于

$$XBEG+(NX-2)\cdot DELX$$

HIS——落入除首尾两个子区间（<*XBEG*和>*XEND*）外的各子区间内的数据个数。

PP——样本χ^2变量。

3.3 子程序说明

子程序PRHIST 为打印直方图的子程序。参数意义同主程序说明。

子程序GCURVE为计算正态分布数值的子程序。其中XM表均值，VAR表方差。其余参数意义同前。

3.4 程序运行及举例

对OUTPUT文件中的ARMA（2,2）过程的200个数据计算其直方图并作正态性检验。

```
C>PDFH  OUTPUT
** TITLE OF THIS SET OF DATA **
ARMA(2,2)
NUMBER OF OBSERVATIONS ---NDATA=?
200
MIN. VALUE FOR ANALYSIS---XBEG=?
INTERVAL BETWEEN CLASSES---DELX=?
NUMBER OF CLASSES          ---NX=?
-10  1.5  15
    ARMA(2,2)
NDATA=    200
 XBEG= -10.000   XEND=       9.500   DELTAX=     1.500
 MEAN=      .009   STANDARD DEVIATION=     3.401

 -10.750   .0
  -9.250  2.0  *X
  -7.750  2.0  XX*
  -6.250  4.0  XXXX*
  -4.750 16.0  XXXXXXXXXXXX*XXX
  -3.250 27.0  XXXXXXXXXXXXXXXXXXXXX*XXXXX
  -1.750 25.0  XXXXXXXXXXXXXXXXXXXXXXXXX     *
   -.250 33.0  XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
   1.250 33.0  XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX*
   2.750 29.0  XXXXXXXXXXXXXXXXXXXXXXXX*XXXX
   4.250 16.0  XXXXXXXXXXXXXXX*
   5.750  8.0  XXXXXXX*
   7.250  3.0  XXX*
   8.750  1.0  *
  10.250  1.0  X

  CHI SQUARE =     7.359
MORE RUN ? (1 FOR YES)
0
```

4. 随机数据的独立性检验（TESTI）

4.1 原理简介

详见第一章1.4节。除了给出统计量Q（见式(1-4-6)），同时给出数据的相关函数及其规范化的值 $\rho(r)/\sqrt{Var\rho(r)}$，其中（见式(1-4-2)）

$$Var(\rho(r)) \approx \frac{1}{N}\left[1+2\sum_{m=1}^{M}\rho^2(m)\right] \qquad r>M$$

4.2 程序功能和参数说明

计算被检验数据的相关函数和规范化相关函数，同时打印相关函数图形。

〔输入量〕

X——数据序列

N——数据个数

KK——用来检验的相关函数值个数

〔输出量〕

R——相关函数值 $\rho(r)$

VAR——$\rho(r)$的方差 $Var\rho(r)$

S——$\rho(r)$ 除以它的标准差（$\sqrt{Var\rho(r)}$）

RKSUM——统计量 $Q=N\sum_{r=1}^{KK}\rho^2(r)$

RR——中间变量

4,3 程序运行及举例

对BEIDT文件中的200个数据序列作独立性检验。

```
C>TYPE  BEIDT
  1.5637    2.0378   -1.4839    .1939     .0240   -2.5525   -1.4129    1.0708
   .9547   -1.6307     .9027   -.1083    1.5958    -.9056   -1.0409    -.1135
  -.7722    1.2029     .3039    .6627     .8193     .5159   -1.1603     .8854
 -1.5423    -.1704     .2648   -.9006     .3956    1.6789     .8027    -.0534
   .0587     .4460    -.0894   2.9030    -.7242    1.5796    -.7115    1.3019
  -.9948    -.6942    -.1671   1.0949     .5024   -1.0040    1.2768     .3687
 -2.1280    1.1979    1.0710   -.1686    -.9868   -2.4023    -.6386     .0474
  1.1089    -.5862    1.3283   1.8631     .5653     .7362    -.9434     .2514
  -.2288    -.2972    1.1468    .9339    2.6635    -.9398    -.4655    1.2750
   .4021     .2910    -.5900  -1.0602   -1.0556    -.1340    -.1844     .3353
  -.4985     .0696   -1.2364  -1.8696    2.9660     .9460   -1.8102    -.7360
  -.2407     .7722   -1.8060   -.0679     .4229    1.3915    1.1430   -1.0995
 -1.1876   -1.3386     .6721   -.5144    -.9527     .4046   -1.6643    1.0111
   .3733    -.2114     .1312   -.3817   -2.3725     .1738     .3695   -1.2384
  -.0479   -1.4403    -.8715   -.6187    -.6461    -.2586   -1.0472   -1.3274
  -.2605     .2400    -.7608   -.6849    -.0170    2.5489    -.1579     .4293
 -1.3913     .1115    -.8439   -.6036     .0590     .0035    -.8822     .5746
 -1.3959    -.5563    3.4267    .5920   -3.1833     .9909    -.4775     .4086
  1.2170    -.6573     .6993    .4715    -.0692    -.2201    -.5657   -1.6322
  2.1246   -1.2467    1.4440   1.0642    -.5428     .7020     .3219   -1.2614
  -.2103    2.0290    -.6499   -.7969     .6988    -.2425    -.5618     .0585
  1.2226     .8120    -.8910   -.4718    -.1167     .5570     .5580   -1.1466
   .3418     .9242     .6325    .1155    1.0894     .5014   -1.6801   -1.6135
 -1.2142     .6456     .4714    .1134    -.8523     .8134    -.4109     .2369
   .4996     .6018     .5329    .9413    -.5030     .2529   -1.8920    -.4348
```

```
C>TESTI  BEIDT
 NO. OF DATA ---  N=? (<=1024)
NO. OF AUTOCORRELATION VALUES USED---  KK=?
200  20
     Q=(SUM OF SQUARES OF AUTOCORRELATIONS)*N= 18.97341
```

LAG K	AUTOCORRELATION	UNIFIED
1	−.0352	−.4980
2	−.1266	−1.7887
3	.0043	.0604
4	−.0276	−.3836
5	−.0067	−.0938
6	.0033	.0452
7	−.0271	−.3769
8	.0420	.5826
9	.1106	1.5327
10	−.0423	−.5792
11	.0295	.4037
12	.0571	.7810
13	.0180	.2447
14	.0428	.5833
15	−.1677	−2.2802
16	.0649	.8604
17	.1213	1.6023
18	−.0796	−1.0382
19	−.0198	−.2565
20	.0127	.1644

```
-.5                    0.0                   .5
.I....I....I....I....I....I....I....I....I....I....I..
                        **
                  *******
                         *
                        **
                         *
                         *
                        **
                         ***
                         ******
                       ***
                         **
                         ***
                         *
                         ***
                *********
                         ****
                         *******
                      ****
                         *
                         *
Stop  —  Program  terminated.
```

5. 随机数据中的隐含周期检验（TESTP）

5.1 原理简介

假设总数为N的数据序列中含有确定性的周期成分，可将其表为

$$x(i)=\frac{a_0}{2}+\sum_{k=1}^{M}\left[a_k\cos(2\pi ki/N)+b_k\sin(2\pi ki/N)\right]$$

其中

$$a_0=\frac{2}{N}\sum_{i=1}^{N}x(i)$$

$$a_k=\frac{2}{N}\sum_{i=1}^{N}x(i)\cos(2\pi ki/N)$$

$$b_k=\frac{2}{N}\sum_{i=1}^{N}x(i)\sin(2\pi ki/N)$$

各谐波振幅为 $c_k=(a_k^2+b_k^2)^{1/2}$ ，其规范化的值定义为

$G_k=c_k^2/\left(\sum_{j=1}^{M}c_j^2\right)$按$G_k$的大小顺序排列以检验振幅的显著性

可以利用第四章4.1.3节介绍的格伦雷德和罗森布特的方法及相应的谱峰显著性表(表4-1)作隐含周期检验，表中数据是按序列数目为奇数（$N=2M+1$）给出的，如果序列总数为偶数，即$N=2M+2$，则省略C_{M+1}而只考虑M个谐波。

5.2 程序功能和参数说明

对具有N个数据的序列给出样本均值，a_0值，以及M个谐波的a_k，b_k，c_k，c_k^2，c_k和$\sum_{j=1}^{M}c_j^2$的值，最后将G_k按大小顺序列表。

当N为奇数时，$M=(N-1)/2$。当N为偶数时$M=N/2-1$

〔输入量〕

N——数据总个数

X——时序数据

〔输出量〕

A——a_k值

B——b_k值

C——c_k值

F——c_k^2值

G——G_k值

5.3 程序运行及举例

对TS文件中的100点数据作周期性检验。

```
C>TYPE  TS
        1.2828           .1378          -.0246          1.1149
        1.6415           .6467          -.5965          -.7772
         .1951           .6151          -.2003         -1.4319
       -1.6639          -.5162           .0140          -.5776
       -1.7932         -1.8201          -.6231           .2210
        -.3810         -1.3392         -1.1519           .1605
        1.0419           .5076          -.4367          -.3207
        1.0246          1.8335          1.1960           .1901
         .1940          1.3692          1.9574          1.1753
         .0505          -.1975           .6347          1.3543
         .4448          -.6994         -1.1211          -.0480
         .4291          -.4614         -1.7634         -1.6578
        -.6287           .0138          -.6894         -1.7592
       -1.6491          -.3886           .3790          -.1294

        -.9795          -.7699           .3602          1.3187
         .7920          -.2992          -.1091          1.2389

        1.9988          1.2246           .2341           .1542
        1.2936          1.7637           .9415          -.2789
        -.4795           .5565          1.0152           .1548
       -1.1362         -1.3190          -.4335           .2353
        -.6729         -1.7861         -1.7938          -.7281
         .0687          -.4834         -1.5394         -1.4491
        -.1698           .7133           .1792          -.6011
        -.5365           .8139          1.5522          1.1031
        -.0144           .0861          1.2830          2.6547
```

C>TESTP TS

NO. OF DATA———N=?

100

SAMPLE MEAN———XM= .503702E—02

A(0)= .100740E—01

	A(K)	B(K)	C(K)	C(K)**2	G
1	.08126	—.04019	.09066	.00822	.00997
2	—.22249	—.27824	.35626	.12692	.15397
3	.51820	.44222	.68124	.46409	.56299
4	.00000	,00000	.00000	.00000	.00000
5	—.02881	.05562	.06264	.00392	.00476
6	—.00421	.02209	.02249	.00051	.00061
7	—.02571	.02555	.03624	.00131	.00159
8	—.00247	.00964	.00995	.00010	.00012
9	—.00934	.03215	.03348	.00112	.00136
10	—.02382	01176	.02656	.00071	.00086
11	—.04000	00000	.04000	.00160	.00194
12	—.01458	—.01369	.02000	.00040	.00249
13	.00000	.00000	.00000	.00000	.00000
14	.00000	.00000	.00000	.00000	.00000
15	.00000	.00000	.00000	.00000	.00000
16	,00000	.00000	.00000	.00000	.00000
17	—.00964	—.01753	.02000	.00040	.00049
18	.03445	—.05429	.06430	.00413	.00502
19	.03583	—.11157	.11718	.01373	.01666
20	.40000	.00000	.40000	.16000	.19410
21	.01990	,07749	.08000	.00640	.00776
22	.01124	.05894	.06000	.00360	.00437
23	.00752	.05953	.06000	.00360	.00437
24	.00126	.01996	.02000	.00040	.00049
25	.00000	.04000	.04000	400160	.00194
26	—.00251	,03992	.04000	.00160	.00194
27	—.00501	,03968	.04000	.00160	.00194
28	—.00750	,03929	04000	.00160	.001[illegible]4
29	—.00995	,03874	.04000	.00160	.00194
30	—.01236	,03804	.04000	.00160	.00194
31	—.01473	,03719	.04000	.00160	.00194

32	—.01703	03619	.04000	.00160	.00194
33	—.02964	.01753	.03443	.00119	.00144
34	—.03072	.01689	.03505	.00123	.00149
35	—.03176	.01618	.03564	.00127	.00154
36	—.03275	.01541	.03619	.00131	.00159
37	—.02000	.00000	.02000	.00040	.00049
38	—.02000	.00000	.02000	.00040	.00049
39	—.02000	.00000	.02000	.00040	.00049
40	—.02000	.00000	.02000	.00040	.00049
41	—.02000	.00000	.02000	.00040	.00049
42	—.02000	.00000	.02000	.00040	.00049
43	—.02000	.00000	.02000	.00040	.00049
44	—.02000	.00000	.02000	.00040	.00049
45	—.02098	—.00618	.02187	.00048	.00058
46	—0.2063	—.00497	.02122	.00045	.00055
47	—.02035	—.00375	.02070	.00043	.00052
48	—.02016	—.00251	.02031	.00041	.00050
49	—.02004	—.00126	.02008	.00040	.00049

SUM OF C(K)**2= .82433

1	.56299	2	.19410	3	.15397	4	.01666	5	.00997
6	.00776	7	.00502	8	.00476	9	.00437	10	.00437
11	.00194	12	.00194	13	.00194	14	.00194	15	.00194
16	.00194	17	.00194	18	.00194	19	.00194	20	.00159
21	.00159	22	.00154	23	.00149	24	.00144	25	.00136
26	.00086	27	.00061	28	.00058	29	.00055	30	.00052
31	.00050	32	.00049	33	.00049	34	.00049	35	.00049
36	.00049	37	.00049	38	.00049	39	.00049	40	.00049
41	.00049	42	.00049	43	.00049	44	.00012	45	.00000
46	.00000	47	.00000	48	.00000	49	.00000		

6. 多项式的提取（POLYTR）

6.1 原理简介

见第一章1.7节。

6.2 程序功能和参数说明

根据选择的多项式阶数（0至4阶），按最小二乘原则提取

数据序列的趋势项，给出多项式本身和提取后的数据序列。

〔输入量〕

K——多项式阶数，取0，1，2，3，4。

N——开始为时序数据总数，若为偶数，则最后改为奇数（舍去最后一个数据）。

X——开始为原始数据序列，最后为提取多项式之后的数据序列。

〔输出量〕

F——$F(1)$至$F(5)$分别表示$\Sigma x(i)$，$\Sigma ix(i)$，$\Sigma i^2x(i)$，$\Sigma i^3x(i)$，$\Sigma i^4x(i)$。

G——$G(1)$至$G(4)$分别为Σi^2，Σi^4，Σi^6，Σi^8的值

C0,C1,C2,C3,C4——分别为多项式系数$c_0,c_1,\cdots,c_4$。

D0,D1,D2,D3,D4——分别为多项式系数$d_0,d_1,\cdots,d_4$。

6.3 程序运行及举例

对TRAN文件中的100点序列用一次多项式拟合。

```
C>TYPE  TRAN
  2.0763   2.0586   1.6647   1.8613   1.9856   2.3614   1.8982   2.1368
  2.1370   2.3908   2.3114   2.9514   2.9057   3.2717   3.2793   3.4300
  3.6665   3.1293   3.1768   3.6977   3.7945   4.0883   3.3791   3.8794
  3.6648   4.2740   4.2610   4.1002   4.6522   4.8112   4.1976   4.4851
  5.1318   4.8250   4.5639   5.1590   5.4785   5.4403   5.7349   5.2470
  6.0675   5.7822   6.0856   5.8740   6.2737   5.9756   5.9910   6.1651
  6.4720   6.1457   6.8260   6.2453   7.0375   6.4172   6.9658   7.2306
  7.3452   6.9154   7.7852   7.2729   7.7701   7.7651   7.6593   7.4702
  7.6872   8.0914   7.9632   7.9566   8.4715   8.0192   9.0716   8.8571
  8.4980   8.6736   9.3598   9.2967   9.1416   9.1795   9.2027   9.4003
  9.7775   9.6621   9.9753  10.2930   9.7805  10.2461  10.0525  10.0997
 10.5257  10.0568  10.8094  10.9455  10.3881  11.2190  10.6211  10.9664
 10.7427  10.8578  11.8624  11.2543
```

```
C>POLYTR TRAN
 NO.  OF DATA---N(<=1024)=?
 ORDER OF POLYNOMIAL---K(<=4)=?
100  1
THE POLYNOMIAL CONTENT IS:
 ( 1.670   )+(   .9844E-01)*t+(   .0000   )*t**2+(   .0000   )*t**3+(   .0000   )*t**4
 DATA AFTER TREND REMOVING:(N=    99)
  .4066880E+00   .2905435E+00  -.2018010E+00  -.1036456E+00  -.7779013E-01
  .1995652E+00  -.3620792E+00  -.2219237E+00  -.3201683E+00  -.1643129E+00
 -.3426575E+00   .1988980E+00   .5475343E-01   .3223088E+00   .2314643E+00
  .2837199E+00   .4217753E+00  -.2138692E+00  -.2648138E+00   .1576416E+00
  .1559972E+00   .3513527E+00  -.4582920E+00  -.5443660E-01   .3674812E+00
  .1432745E+00   .3182989E-01  -.2274146E+00   .2261408E+00   .2866962E+00
 -.4253486E+00  -.2362932E+00   .3119626E+00  -.9328232E-01  -.4528267E+00
  .4382868E+01   .2648841E+00   .1282396E+00   .3243951E+00  -.2619492E+00
  .4601061E+00   .7636131E-01   .2813168E+00  -.2872758E-01   .2725280E+00
 -.1240170E+00  -.2070611E+00  -.1314058E+00   .7704971E-01  -.3476950E+00
  .2341607E+00  -.4449843E+00   .2487713E+00  -4699731E+00  -.1981791E-01
  .1555379E+00   .1626932E+00  -.3655514E+00   .4058042E+00  -2049404E+00
  .1938151E+00   .9037042E-01  -.1138743E+00  -.4014186E+00  -.2828632E+00
  .2289235E-01  -.2037522E+00  -.3087967E+00   .1076589E+00  -.4430857E+00
  .5108694E+00   .1979244E+00  -.2596195E+00  -.1824640E+00   .4052916E+00
  .2437462E+00  -.9798206E-02  -.7034279E-01  -.1455873E+00  -.4643149E-01
```

```
 .2323241 E+00    .1847922 E-01    .2332347 E+00    .4524905 E+00   -.1584539 E+00
 .2087016 E+00   -.8334361 E-01   -.1345880 E+00    .1929671 E+00   -.3743772 E+00
 .2797780 E+00    .3174342 E+00   -.3384110 E+00    .3940446 E+00   -.3022994 E+00
-.6644388 E-01   -.3775893 E+00   -.3609339 E+00    .5452220 E+00
Stop - Program terminated.
```

7. 逆序检验（INORD）

7.1 原理简介

见第一章1.5节中的式（1-5-5）至（1-5-9）及相应说明。

7.2 主程序功能和参数说明

将数据序列分成 K 段，计算各段的样本均值，并给出均值非平稳性的逆序检验统计量 U。

〔输入量〕

X——时序数据

N——时序数据总数

K——分段数目

〔输出量〕

XA——各段的样本均值

U——逆序检验的统计量。

7.3 子程序说明

子程序INVT用于计算被检验序列的逆序统计量U，程序中X为被检验的序列（在本主程序中为均值序列），K为被检验序列的个数。

7.4 程序运行及举例

对文件TRAN中的序列的均值分成十段，进行均值的逆序检验

```
C>INORD  TRAN
NO. OF  DATA————N=?
NO. OF  SEGMENTS————K=?
100  10
  U=  4.1144
  (FOR A STATIONARY SERIES, ABS(U)  <= 2)
Stop — Program terminated.
```

8. 游程检验（RUNTES）

8.1 原理简介

见第一章1.5节

8.2 主程序功能和参数说明

对观察序列$\{x_t\}$及给定门限值，用游程检验方法计算μ_r,σ_r

及检验统计量 $z=\frac{r-\mu_r}{\sigma_r}$

〔输入量〕

X——观察序列$\{x_t\}$。

E——给定门限值。当$x_t \geq E$时取值为1；反之则为0。

〔输出量〕

N1——$\{x_t\}$中小于门限值的数据个数。

N2——$\{x_t\}$中大于门限值的数据个数。

IR——$\{x_t\}$中含游程个数r。

E——均值估计μ_r。

SIGM——方差估计σ_r

Z——检验统计量 $z=\frac{r-\mu_r}{\sigma_r}$

8.3 程序运行及举例

建立数据文件RUNTES.O

```
1.,1.,0.,0.,0.,1.,0.,0.,1.,1.,
1.,1.,1.,0.,0.,1.,0.,0.,1.,1.,
0.,1.
```

运行结果

```
C>RUNTES RUNTES.O
  NUMBER OF THE DATA  N=?
22
  THRESHOLD  E=?
0.
  N1= 10    N2=12    IR= 11
  AVERAGE= 11.91      VARIANCE= 2.269
    Z=-.4007
Stop — Program terminated.
```

9. 格林函数的计算（GREENF）

9.1 原理简介

格林函数的递推计算公式见式（2-3-68）。

本程序使用的符号是φ_i和θ_i（和a_i，b_i反号）。

9.2 主程序功能和参数说明

对参数已知的ARMA 序列，计算其格林函数。

〔输入量〕

IP，IQ——序列的AR-及MA-阶数（n和m）

F，C——序列的AR及MA参数（φ_i和θ_i）

L——欲求格林函数的步数。

〔输出量〕

G——序列的格林函数G_1，G_2，…，G_L。

9.3 程序运行及举例。

对$\varphi_1=0.5$，$\theta_1=0.5$的ARMA（1,1）模型，运行本程序的结果是

```
C>GREENF
AR- AND MA- ORDER OF MODEL IP=? , IQ=?
1, 1
AR--PARAMETERS  F(I), I=1,...,1
.5
MA--PARAMETERS  C(I),I=1,...,1
```

.8
STEPS OF GREEN FUNCTION L=?
8

```
ARMA(1,1)

              F=     .500
              C=     .800

      THE GREEN FUNCTION IS:

   G(0)=       1.0000
   G(1)=    -.3000E+00
   G(2)=    -.1500E+00
   G(3)=    -.7500E-01
   G(4)=    -.3750E-01
   G(5)=    -.1875E-01
   G(6)=    -.9375E-02
   G(7)=    -.4688E-02
   G(8)=    -.2344E-02
Stop — Program terminated.
```

10. 逆函数的计算（INF）

10.1 原理简介

逆函数的递推计算公式见式（2-3-79），但本程序使用的符号是 φ_i 和 θ_i（和 a_i，b_i 反号）。

10.2 主程序功能和参数说明

对参数已知的ARMA序列，计算其逆函数。

〔输入量〕

IP，IQ——序列的AR-及MA-阶数（n和m）。

F，C——序列的AR及MA参数（φ_i和θ_i）。

L——欲求逆函数的步数。

〔输出量〕

I——序列的逆函数 I_1，I_2，$\cdots I_L$。

10.3 程序运行及举例

对$\varphi_1=0.5$，$\theta_1=0.8$的ARMA（1,1）模型，运行本程序的结果是

```
C>INF
AR— AND MA— ORDER OF MODEL IP=? , IQ=?
1,1
AR—PARAMETERS F(I), I=1,...,1
.5
MA—PARAMETERS C(I),I=1,...,1
.8
STEPS OF INVERSE FUNCTION   L=?
8
   ARMA(1,1)

                 F=    .500
                 C=    .800

        THE INVERSE FUNCTION IS:

   I(0)=     -.1000E+01
   I(1)=     -.3000E+00
   I(2)=     -.2400E+00
   I(3)=     -.1920E+00
   I(4)=     -.1536E+00
   I(5)=     -.1229E+00
   I(6)=     -.9830E-01
   I(7)=     -.7864E-01
   I(8)=     -.6291E-01
 Stop — Program terminated.
```

11. 朱利准则（JURYC）

11.1 原理简介

见第二章2.3节

11.2 主程序功能和参数说明

对给定模型的 AR（MA）参数，检验其是否在平稳(可逆）域内，即相应特征方程的根是否都在 B平面的单位圆外。

〔输入量〕

N——模型的AR或MA阶数

F——模型的AR参数$\varphi_1, \varphi_2, \cdots, \varphi_N$或$MA$参数$\theta_1, \theta_2, \cdots \theta_N$。

〔输出量〕

JURYC——出口标记。JURYC＝0，特征方程根都在单位圆外，相应模型在平稳（或可逆）域内；JURYC＝1 则相反。

11.3 程序运行及举例

对参数为 $\varphi_1=-0.8$，$\varphi_2=-0.5$的 AR(2)模型，检验该模型是否平稳。程序运行结果是

```
C>JURYC
 AR/MA — ORDER  OF  MODEL  N=?
2
 AR/MA — PARAMETERS  OF  MODEL  F(I),I=1,…,2 ?
-.8,-.5
      JURYC= 0
 ALL  ROOTS  ARE  OUTSIDE  THE  UNIT  CIRCLE(STABLE)
Stop — Program  terminated.
```

12. 理论自相关函数的计算（AUTOCR）

12.1 原理简介

设ARMA(n,m)模型的参数为$\varphi_1, \cdots, \varphi_n$；$\theta_1, \cdots, \theta_m$，$\{\rho_i\}$为其相关函数序列。由第二章 2.3 节式（2-3-72）可得下面线性方程组（令$l=max(n,m+1)$）

$$\rho(k)-\varphi_1{}^*\rho(k-1)-\cdots-\varphi_l{}^*\rho(k-l)=\sigma_a^2\sum_{j=k}^{l}\theta_j^* G_{j-k}$$

$$k=0,1,2,\cdots,l-1$$

其中σ_ε^2是模型白噪声方差，G_j是相应的格林函数。

$$\varphi_j^*=\begin{cases}\varphi_j & 若\ 1\leq j\leq n\\ 0 & 若\ j>n\end{cases}$$

$$\theta_j^*=\begin{cases}\theta_j & 若\ 1\leq j\leq m\\ 0 & 若\ j>m\end{cases}$$

由上述线性方程组可将$\rho(0),\rho(1),\cdots,\rho(l-1)$求出。对于$k\geq l$，$\rho(k)$可按如下递推公式计算

$$\rho(k)=\varphi_1\rho(k-1)+\varphi_2\rho(k-2)+\cdots+\varphi_n\rho(k-n)\quad k\geq l$$

12.2 主程序功能和参数说明

对阶数及参数已知的 ARMA 模型求其理论自相关函数。

〔输入量〕

IP,IQ——模型的AR- 及MA- 阶数（n和m）

F，C——模型的AR参数$\varphi_1,\cdots,\varphi_{IP}$及$MA$参数$\theta_1,\cdots,\theta_{IQ}$。

M——欲计算自相关函数的步数。

SIGM——模型白噪声的方差（σ_ε^2）

〔输出量〕

RO——序列方差的理论值。

RR——理论自相关函数$\rho(1),\cdots,\rho(M)$

ISW——出口标记。ISW＝0是正常出口；$ISW=1$是正非常出口，线性方程组退化无解。

12.3 子程序说明

12.3.1 子程序GREENF

见本附录9。

12.3.2 子程序GAUSS

功能：同主元素消去法求解n阶线性方程组

$$AX=B$$

其中A是$n\times n$方阵，X，B是n维向量。

〔输入量〕

N——线性方程组阶数。

A——线性方程组左端系数矩阵。

B——线性方程组右端系数向量。

EPS——机器零。

〔输出量〕

B——线性方程组的解向量。

ISW——出口标记。ISW＝0是正常出口；ISW＝1是非正常出口，线性方程组退化无解。

12.4　程序运行及举例

对参数为$\varphi_1=0.5$，$\theta_1=0.8$的ARMA(1,1)模型，调用本程序计算理论相关函数，运行结果是

```
C>AUTOCR
 AR- AND MA- ORDER OF THE MODEL   IP=? , IQ=?
1, 1
 AR- PARAMETERS OF MODEL  F(I), I=1,…, IP  ,
.5
 MA- PARAMETERS OF MODEL C(I),I=1,…,IQ  ,
.8
 MEAN SQUARE OF THE SEQUENCE  SIGM=?
1.
 STEPS OF CORRELATION  FUNCTION  M=?
8
 PARAMETERS  OF  ARMA(1,1)
          F=        .50
          C=        .80
        VARIANCE  OF  THE MODEL RO=     1.12000
 CORRELATION VALUES  OF  THE  MODEL  ARE:
   RR(1)=        -.21429
   RR(2)=        -.10714
```

```
    RR(3)=          -.05357
    RR(4)=          -.02679
    RR(5)=          -.01339
    RR(6)=          -.00670
    RR(7)=          -.00335
    RR(8)=          -.00167
Stop — Program terminated.
```

13. 偏相关函数的递推计算（PARCR）

13.1 原理简介

偏相关函数的递推计算公式见第三章3.5节式（3-5-8）

13.2 主程序功能和参数说明

由已求出的自相关函数递推计算序列的偏相关函数。

〔输入量〕

M——欲求偏相关函数的阶数。

RR——已求出的自相关函数 $\rho(1),\rho(2),\cdots,\rho(M)$

〔输出量〕

FF——M阶偏相关函数φ_{M1}，$\varphi_{M2},\cdots,\varphi_{MM}$

FFK——存放φ_{11}，$\varphi_{22},\cdots,\varphi_{MM}$

13.3 程序运行及举例

用已得到的自相关函数建立数据文件PARCR.O

-.53333，-.07333，

运行输出打印结果中，按照顺序将φ_{11}；$\varphi_{21},\varphi_{22},\cdots,\varphi_{M1},\varphi_{M2},\cdots,\varphi_{MM}$ 打印，并将 φ_{11}，$\varphi_{22},\cdots,\varphi_{MM}$ 输出。调用本程序运行结果是

```
C>PARCR  PARCR.o
  NUMBER  OF  AUTO— AND
PARTIAL  CORRELATION  FUNCTIONS  M=
2
        K=      1
```

```
         -.5333
   K=      2
         -.8000      -.5000
 2-ORDER  PARTIAL  CORRELATION  IS:
   FF(1,1)=          -.5333
   FF(2,2)=          -.5000
Stop - Program  terminated.
```

14. ARMA谱的计算和作图（ARMASP）

14.1 原理简介

ARMA 模型$x_t-\phi_1x_{t-1}-\cdots-\phi_nx_{t-n}=\varepsilon_t-\theta_1\varepsilon_{t-1}-\cdots-\theta_m\varepsilon_{t-m}$的系数$\phi_i$，$\theta_i$及$\varepsilon_t$和$x_t$的方差（$\sigma_\varepsilon^2$和$\sigma_x^2$）已知，可按下式计算$x_t$的单边归一化功率谱：

$$S_x(f)=\frac{2\sigma_\varepsilon^2\Delta}{\sigma_x^2}\left|\frac{1-\theta_1B-\cdots-\theta_mB^m}{1-\phi_1B-\cdots-\phi_nB^n}\right|^2_{B=\exp(-j2\pi f\Delta)}$$

$$0\leq f\leqslant 1/(2\Delta)$$

其中Δ为数据的采样间隔。

14.2 主程序功能和参数说明

输入 ARMA 模型的已知参数，并选定频率 f 的变化范围和在该范围内的取值点数，本程序计算出相应功率谱 $S_x(f)$ 的分贝值，即$20\log_{10}S_x(f)$。f 的最小值（下限）为 0，最大值(上限)为 $1/(2\Delta)$，若选定的最大频率超过上限时自动由程序改为上限值。

〔输入量〕

NAR——自回归阶数n

NMA——滑动平均阶数m

PHI——自回归系数$\phi_1,\cdots,\phi_n$。

THETA——滑动平均系数$\theta_1,\cdots,\theta_m$

XMIN——要求计算的最低频率

XMAX——要求计算的最高频率

N——XMIN和XMAX之间所取的点数

GAMAZ——x_t的方差σ_x^2

SIGZ——ε_t的方差σ_ε^2

DEL——数据采样间隔Δ。

〔输出量〕

X——频率值f

YS——$S_x(f)$的分贝值

RUP，IUP分别为$1-\theta_1 B-\cdots-\theta_m B^m |_{B=\exp(-j2\pi f\Delta)}$ 的实部和虚部。

RDWN，IDWN分别为$1-\phi_1 B-\cdots-\phi_n B^n |_{B=\exp(-j2\pi f\Delta)}$的实部和虚部。

14.3　子程序说明

子程序GRAPH用于单变量函数的打点作图。N为点数，X为纵坐标。

14.4　程序运行及举例

已知AR(2)模型的参数——$\varphi_1=1.03$，$\varphi_2=-0.6$，$\sigma_\varepsilon^2=1$，$\sigma_x^2=1$，$\Delta=1$，在频率为0至0.5之间计算并绘出功率谱密度曲线（30点）。

```
C>ARMASP
 AR  ORDER--- n = ?
 MA  ORDER --- m = ?
 MIN. FREQ. OF PSD---Fmin=?
 MAX. FREQ. OF PSD---Fmax=?
 POINTS OF PSD BETWEEN Fmin AND Fmax --- N=?
2  0  0  .5  30
 VARIANCE OF Xt = ?
 VARIANCE OF RESIDUAL Et = ?
 SAMPLING INTERVAL---DEL=?
```

1 1 1

AR PARAMETERS———PH1, PH2,…,PHn = ?

1.03 —.6

NORMALIZED LOG POWER SPECTRAL DENSITY(db)

FREQ	DB	FREQ	DB
.000	15.7856	.017	16.0223
.034	16.7413	.052	17.9669
.069	19.7266	.086	22.0042
.103	24.5821	.121	26.6523
.138	26.6531	.155	23.9403
.172	19.8450	.190	15.6574
.207	11.8539	.224	8.5138
.24.	5.6010	.259	3.0586
.276	.8337	.293	—1.1169
.310	—2.8280	.328	—4.3269
.345	—5.6352	.362	—6.7702
.379	—7.7456	.397	—8.5722
.414	—9.2587	.431	—9.8120
.448	—10.2372	.466	—10.5383
.483	—10.7179	.500	—10.7776

PLOT PSD (1 =YES) ?

1

```
                                    ****MAX Y=  .2665315E+02
 .1579560E+02                              .                        *
 .1602230E+02                              .                        *
 .1674127E+02                              .                         *
 .1796688E+02                              .                           *
 .1972682E+02                              .                              *
 .2200420E+02                              .                                 *
 .2458213E+02                              .                                     *
 .2665229E+02                              .                                        *
 .2665315E+02                              .                                        *
 .2394031E+02                              .                                    *
 .1984499E+02                              .                              *
 .1565743E+02                              .                        *
 .1185392E+02                              .                  *
 .8513765E+01                              .             *
 .5601011E+01                              .         *
 .3056538E+01                              .     *
 .8337398E+00                              . *
-.1116911E+01                            * .
-.2827993E+01                          *   .
-.4326883E+01                        *     .
-.5635223E+01                      *       .
-.6770220E+01                    *         .
-.7745567E+01                   *          .
-.8572184E+01                 *            .
-.9258721E+01                *             .
-.9811996E+01                *             .
-.1023724E+02               *              .
-.1053832E+02               *              .
-.1071793E+02              *               .
-.1077763E+02              *               .
Stop - Program terminated.
```

15. 样本自相关和互相关的快速计算（CORELA）

15.1 原理简介

见附录一第6节

15.2 主程序功能和参数说明

计算序列x_t的自相关或长度相同的两个序列（x_t和y_t）的互相关。是否将序列的均值扣除及是否打印相关函数图可由用户决定。

〔输入量〕

XD——时序数据x_t

YD——时序数据y_t

N——时序x_t或y_t的数据个数

M——分段后每段的数据个数

MODE——算法标志，0为不扣除均值的自相关；1为不扣除均值的互相关；2为扣除均值的自相关；3为扣除均值的互相关。

IE——是否打印相关函数图标志，1为打印，0为不打印。

〔输出量〕

ILAG——存贮相关函数的迟后值

XA——开始存贮分段的x_t和y_t数据，最后存贮相关函数值。

XC，XS，ZC，ZS为中间变量。

15.3 子程序说明

子程序FFT对二实数序列同时求快速富氏正变换或反变换。INN＝0时作正变换；INV＝1时作反变换。XC和XS分别为二实数数组，表示某一复数数组的实部和虚部，开始为原输入数组，最后为变换后的数组。N为数组（XC或XS）所含的序列长度，为2的某次幂。

子程序GETD将一个长的实数序列XD分段读取（每调用一次读取一段），构成序列X。NRD为读取序列的长度。IS是每次

读取的开始点在长序列XD中的位置。

子程序GRAPH见程序 ARMASP 说明。

15.4 程序运行及举例

计算AR(1)过程(数据文件AR1) 和与其滞后 8 步的同一过程(数据文件AR1D)之间的互相关。

```
C>CORELA AR1D AR1
SECTION SIZE---M=
64
TOTAL NO. OF DATA---N=
512
MODE=
            MODE = 0 AUTO CORRELATION WITHOUT MEANS
                     SUBTRACTED
            MODE = 1 CROSS CORRELATION WITHOUT MEANS
                     SUBTRACTED
            MODE = 2 AUTO CORRELATION WITH MEANS
                     SUBTRACTED
            MODE = 3 CROSS CORRELATION WITH MEANS
                     SUBTRACTED
3
M=  64  N=  512  MODE=3
XMEAN=     -.14826E+00    YMEAN=     -.14360E+00
PLOT  THE  CORRELATION  FUNCTION  (1=YES, 0=NO)=
1
```

```
***MAX=  .1458992E+01

-.5479182E-01
-.5844686E-01
-.4798355E-01
 .3415444E-01
 .1372937E+00
 .2385560E+00
 .4119888E+00
 .8175052E+00
 .1458992E+01
 .8164825E+00
 .4066499E+00
 .2294864E+00
 .1297970E+00
 .3173815E-01
-.4545777E-01
-.6116499E-01
-.4676609E-01
 .1036205E-01
-.4404825E-01
-.4239894E-01
-.2230968E-01
-.1661919E-01
 .2255607E-01
 .2961340E-01
 .1408034E+00
 .1290922E+00
 .9493501E-01
 .1067802E+00
 .1191620E+00
 .7812975E-01
 .3472123E-01
-.1603111E-02
-.7719047E-01
```

CORRELATION FUNCTION

LAG	CORR	LAG	CORR	LAG	CORR
0	-.548E-01	1	-.584E-01	2	-.480E-01
3	.342E-01	4	.137E+00		
5	.239E+00	6	.412E+00	7	.818E+00
8	.146E+01	9	.816E+00		
10	.407E+00	11	.229E+00	12	.130E+00
13	.317E-01	14	-.455E-01		
15	-.612E-01	16	-.468E-01	17	.104E-01
18	-.440E-01	19	-.424E-01		
20	-.223E-01	21	-.166E-01	22	.226E-01
23	.296E-01	24	.141E+00		
25	.129E+00	26	.949E-01	27	.107E+00
28	.119E+00	29	.781E-01		
30	.347E-01	31	-.160E-02	32	-.772E-01

Stop - Program terminated.

16. 逆函数法估计模型参数的初值（INGUS）

16.1 原理简介

参数估计的逆函数方法见第三章 3.6 节

16.2 主程序功能和参数说明

对于阶数给定的 ARMA 模型，由已求出的自相关函数，用逆函数法计算模型参数的估计值。如MA参数不在可逆域内，则应将其调整到可逆域内。

〔输入量〕

IP，IQ——模型的AR及MA参数。

R——已得出的自相关函数值$\rho_1,\cdots,\rho_{IPST}$，

$$IPST=\max(IP,IQ)+IQ$$

〔输出量〕

F——AR参数的估计值。

C——MA参数的估计值，在可逆域内。

16.3 子程序说明

16.3.1 子程序ARMAI

功能：由相关函数用逆函数法计算阶数为给定模型的参数估计值。

〔输入量〕

IP,IQ,R——见主程序参数说明

〔输出量〕

F,C模型*AR*及*MA*参数的估计值，*MA*参数不一定在可逆域内。

16.3.2 子程序PARCR

见本附录13

16.3.3 子程序GAUSS

见本附录12.3.2

16.3.4 子程序ROOT

功能：对给定多项式，将其（*B*平面）单位圆内的根置换到单位圆外，并求出变换后的多项式

〔输入量〕

MM——多项式的阶数。

A——形如$1-\varphi_1B-\varphi_2B^2-\cdots-\varphi_{MM}B^{MM}$多项式的系数$1,\varphi_1,\varphi_2,\cdots,\varphi_{MM}$

〔输出量〕

A——经变换后多项式$1-\varphi_1'B-\varphi_2'B^2-\cdots-\varphi_{MM}'B^{MM}$的系数$1,\varphi_1',\varphi_2',\cdots,\varphi_{MM}'$。该多项式的所有根都在单位圆外。

16.3 子程序RTPOL

功能：对给定多项式求其全部根及各根的模。

〔输入量〕

N——多项式阶数

A——多项式$a_1Z^N+a_2Z^{N-1}+\cdots+a_{N+1}$的系数$a_1, a_2 \cdots a_N$。

L——允许的最大迭代次数。

K——根的有效位数。

〔输出量〕

B——依次存放各根的实部。

C——依次存放各根的虚部。

16.4 程序运行及举例

对于MA(2)模型，用已得出的自相关函数建立数据文件INGUS.O

.002, .3, .04, .6,

用逆函数法进行参数估计，调用本程序的运行结果是：

```
C>INGUS  INGUS.o
 AR— AND MA— ORDER OF THE NODEL  IP=?  IQ=?
0,2
 MA— PARAMETERS:
      4.9322        2.4269
    REAL—PART    IMAGINARY—PART    MAGNITUDE
   OF THE ROOT    OF THE ROOT    OF THE ROOT
       .18577          .00000          .18577
     —2.21811          .00000         2.21811
*****************************
    REAL—PART    IMAGINARY—PART    MAGNITUDE
   OF THE ROOT    OF THE ROOT    OF THE ROOT
     —2.21811          .00000         2.21811
      5.38305          .00000         5.38305
 $$$ ESTIMATION OF MA— PARAMETERS  ARE:
   —.265    .084
Stop — Program terminated.
```

17. ARMA模型参数的矩估计（ARMAME）

17.1 方法简介

ARMA模型参数的矩估计方法见第三章3.2节。

17.2　主程序功能和参数说明

对阶数为已知的 ARMA 模型，由相关函数计算模型参数的估计值。

〔输入量〕

IP,IQ——模型的AR-及MA-阶数。（n和m）

RR——已求出的自相关函数。

ML——允许的最大迭代次数。

EPS1——迭代精度

EPS2——计算机上最小浮点数。

〔输出量〕

,C——模型AR及MA参数的矩估计（ϕ_i和θ_i）

17.3　子程序说明

17.3.1　子程序MAME

功能：对于阶数为已知的MA模型，由相关函数计算模型MA参数的估计值。

〔输入量〕

IQ——模型的MA阶数。

RO——序列的方差

RR——已得到的自相关函数。

EPS——迭代精度。

ML——允许的最大迭代次数。

SIGM——白噪声方差。

〔输出量〕

C——模型MA参数的矩估计值。

17.3.2　子程序GAUSS

见12.3.2

17.4　程序运行及举例

对于ARMA（3,1）模型，用已得到的自相关函数建立数据

文件ARMAME.O

—.70175,.27193,—.16667,.20789,

求该模型参数的矩估计，运行结果是

```
C>ARMAME ARMAME.0
 AR— AND MA— ORDER OF THE MODEL  IP=?   IQ=?
3,1
 MAXIMUM NUMBER  OF  ITERATION  ML=?
10
 TOLERANCE VALUE  EPSI=?
.001
ARMA(3,1) PARAMETERS  ESTIMATION:
   —.800          —.500          —.300
   .6997E+00
Stop — Program terminated.
```

18．线性最小二乘法(LSME)

18.1 方法简介

线性最小二乘参数估计方法见第三章3.1节及第八章8.2.1节

18.2 程序功能和参数说明

本程序是一个独立模块，用于求已知形式模型的线性最小二乘估计。子程序SUBLSM由用户自行编制。该程序运行实例见27,28,29。

〔输入量〕

MRE——线性模型的参数个数。

N1,N2——建立线性模型时，对观察序列使用样本的起始及终止标号。

〔输出量〕

F0,F——模型参数的最小二乘估计。F0是常数项。

S0——模型残差的标准差。

AIC——模型的AIC值。

ISW——出口标记。$ISW=0$是正常出口；$ISW=1$是非正常出口，对称方程组退化无解。

18.3 子程序说明

18.3.1 子程序SUBLSM

功能：根据具体线性模型构造a_{ij}及z_i。本子程序由使用者自行编制。

〔输入量〕

N——每一变元观察数据的个数。

IS——观察数据的变元个数。

X——观察数据序列$\{X_t\}$。

MRE——模型参数个数。

K——$K\in[N1,N2]$，$N1,N2$含义见程序LSM。

NO——实际使用数据比$N1\sim N2$减少的个数。

〔输出量〕

A——存放a_{ij}，$1\leqslant j\leqslant MRE$，

Zk——存放z_i。

18.3.2 子程序LDLT

功能：用平方根法解对称正定线性方程组。

$$AX=B$$

其中A是N×N矩阵，X,B是N维向量

〔输入量〕

A——对称正定矩阵，是线性方程组左端系数矩阵

B——方程组右端的系数向量

N——线性方程组阶数

〔输出量〕

X——线性方程组的解向量

ISW——出口标记。$ISW=0$是正常出口；ISW=1是非正常出口，线性方程组退化无解。

19. *Marquardt*阻尼最小二乘法(MARQT)

19.1 原理简介

为了用ARMA模型拟合平稳随机过程，采用非线性最小二乘方法估计模型的参数。

首先将过程x_t作规范化处理（减去均值，再除以标准差），根据预定阶数（n和m）和系数的初始估计（ϕ_i和θ_i的初值），按以下方法进行多次迭代运算对参数作精估计

令$\beta=(\phi_1,\cdots\phi_n,\theta_1,\cdots,\theta_m,\mu)^T$表参数向量($\mu$为$x_t$的均值)，则第$k$次和第$k+1$次递推结果的关系为

$$\beta^{(k+1)}-\beta^{(k)}=A^{(k)}X^{(k)^T}e^{(k)}$$

其中

$$A^{(k)}=(X^{(k)^T}X^{(k)}+\lambda_k I)^{-1}$$

$X^{(k)}$和$e^{(k)}$是第k次迭代所应的雅可比阵和残差向量。这里雅可比阵是指

$$X=\begin{pmatrix} \frac{\partial f_1}{\partial\phi_1} & \frac{\partial f_1}{\partial\phi_2} \cdots \frac{\partial f_1}{\partial\phi_n} & \frac{\partial f_1}{\partial\theta_1} \cdots \frac{\partial f_1}{\partial\theta_m} & \frac{\partial f_1}{\partial\mu} \\ \frac{\partial f_2}{\partial\phi_1} & \frac{\partial f_2}{\partial\phi_2} \cdots \frac{\partial f_2}{\partial\phi_n} & \frac{\partial f_2}{\partial\theta_1} \cdots \frac{\partial f_2}{\partial\theta_m} & \frac{\partial f_2}{\partial\mu} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_N}{\partial\phi_1} & \frac{\partial f_N}{\partial\phi_2} \cdots \frac{\partial f_N}{\partial\phi_n} & \frac{\partial f_N}{\partial\theta_1} \cdots \frac{\partial f_N}{\partial\theta_m} & \frac{\partial f_N}{\partial\mu} \end{pmatrix}$$

其中 $f_t=\mu+\phi_1(x_{t-1}-\mu)+\cdots+\phi_n(x_{t-n}-\mu)-\theta_1\varepsilon_{t-1}-\cdots$
$\cdots-\theta_m\varepsilon_{t-m}$

$t=1,2,\cdots,N$

残差向量是指

$\mathbf{e}=(\varepsilon_1,\cdots,\varepsilon_N)^T$

$\lambda^{(k)}$是第k次迭代的阻尼系数。

结束迭代的条件为以下三者之一满足，即

i) 迭代次数超过预定值$ITMAX$。

ii) 前后两次迭代结果的全部参数满足

$$\left|\frac{\boldsymbol{\beta}^{(k+1)}-\boldsymbol{\beta}^{(k)}}{\boldsymbol{\beta}^{(k+1)}}\right|<\text{某给定值TOL(1)}$$

iii) 前后两次拟合残差平方和的相对改善满足

$$\left|\frac{\text{SSNEW}-\text{SSOLD}}{\text{SSNEW}}\right|<\text{某给定值TOL(2)}$$

或连续三次迭代而无改善（SSNEW=SSOLD）。

其中SSOLD和SSNEW分别为第k次和第$k+1$次迭代结果所对应的残差平方和。

最终的参数估计值连同它们的95%置信限为

$$\boldsymbol{\beta}^{(k+1)} \pm 2\sigma_{\varepsilon}\sqrt{\mathrm{diag}(A^{(k)}X^{(k)T}X^{(k)}A^{(k)})}$$

其中σ_ε为残差的标准差。

残差平方和对于各估计参数的梯度的模为$2X^T e$。为了模型定阶的参考，可按式(3-5-34)和（3-5-35）计算AIC和BIC值。

19.2　主程序功能和参数说明

对设定阶数(n，m）和系数(ϕ_i，θ_i)初值的过程(均值μ的初值是由程序求得的样本平均）用非线性最小二乘方法作精估计。均值是否作为精估计的参数之一由用户决定。

最大迭代次数及TOL(1)，TOL(2)二给定值也由用户选择。程序执行过程中每次迭代的阻尼比（λ值）和残差平方和均在屏幕显示。

拟合结果给出X_t的方差($\Sigma(X-\mu)^2$/N)和平均值(ΣX/N）。如果要求对均值作精估计，则给出

$$\begin{array}{l} \qquad\quad \mu+(95\%\text{置信限}) \\ \text{MEAN}=\mu \\ \qquad\quad \mu-(95\%\text{置信限}) \end{array}$$

ϕ_i和θ_i的估计结果，也以和μ的估计相类似的方式给出。$\phi_1,\cdots,\phi_n$和$\theta_1,\cdots,\theta_m$均自左至右顺序排列。最后一个估计参数是σ_ε^2(SIGMA　E—SQUARE)，它是指$\sum_{t=1}^{N}\varepsilon_t^2/N$。

残差平方和对参数的梯度之模自左至右依次为对于$\phi_1,\cdots,\phi_n,\theta_1,\cdots,\theta_m,(\mu)$而言。

程序给出了拟合结果的AIC和BIC值。

〔输入量〕

N(或NAR)——自回归阶数。

M(或NMA）——滑动平均阶数。

NOB——x_t数据个数。

X——开始存放x_t的数据值

ITMAX——规定的最大迭代次数

TOL(1)——规定的参数相对变化的允许值

TOL(2)——规定的残差平方和相对变化的允许值。

IMEAN——是否估计均值标志，$IMEAN=0$表不估计均值。

PAR——开始依次为$\phi_1,\cdots,\phi_n,\theta_1,\cdots,\theta_m$的初值，最后则依次为它们的估计值。

〔输出量〕

OBS——规范化的时序数据值。

AT——残差ε_t序列。

NPAR——待估参数总数(IMEAN=0时NPAR=$N+M$，否则NPAR=$N+M+1$)。

BNDLW——（参数估值）$-2\times$（估计参数的标准差）。

BNDUP——（参数估值）$+2\times$（估计参数的标准差）。

GRD——残差平方和对参数的梯度之模(依次为对于$\phi_1\cdots$, $\phi_n,\theta_1,\cdots,\theta_m$和$\mu$而言)。

SIGMA——最后表示(实际残差平方和)/(时序数据个数)。

VAR——时序数据的样本方差值。

19.3 子程序说明

本程序共调用了六个子程序

子程序MAR是麦夸特阻尼最小二乘法估计程序，根据在19.1中给出的迭代公式计算每次迭代的参数修正量$\delta\boldsymbol{\beta}$(即$\boldsymbol{\beta}^{(k+1)}-\boldsymbol{\beta}^{(k)}$)，开始时取$\lambda=1.5$(NPAR)，若迭代结果新的残差平方和更大或无改善(SSNEW≥SSOLD)则增大λ($\lambda^{(k+1)}=4\lambda^{(k)}$)，反之则减小$\lambda$($\lambda^{(k+1)}=\lambda^{(k)}/1.8$)。

*MAR*中的参量有：

S—X^Te；DX—$\delta\boldsymbol{\beta}$；A—X^TX中各元素；

XTX——未加入阻尼项的雅可比矩阵元素；

ZA，ZB，ZC——分别为$\partial f_i/\partial \phi_1$，$\partial f_i/\partial \theta_1$，$\partial f_i/\partial \mu$ ($i=1,\cdots,N$)；

D——2×(估计参数的标准差)；

SIGMA——（规范化的残差平方和）/NOB。

PAR意义同主程序参数

子程序CAT用于计算残差ε_t，有关参数意义同前。

子程序EQN用作线性方程组求解，即由$\boldsymbol{XB}=\boldsymbol{Y}$求解$\boldsymbol{B}$，其中$\boldsymbol{Y}$和$\boldsymbol{B}$为$M\times1$阵，$\boldsymbol{X}$为$M\times M$阵，$M$是方阵$\mathbf{X}$的维数。

子程序MATINV用于矩阵求逆，其中A为方阵，N为其维数。

子程序JAC用来计算雅可比阵的一列元素。JCOL是根据待估参数决定雅可比阵形式的标志数。THETA为$\theta_1,\cdots,\theta_m$，W为中间变量，PAR意义同前。

子程序CI用于计算95%置信区间，其中

A——表示$\boldsymbol{A}^{(k)}$

B——表示$\boldsymbol{X}^{(k)T}\boldsymbol{X}^{(k)}$

C——A和B的乘积

D——2×（估计参数的标准差）

19.4 程序运行及举例

```
C>TYPE D70
47., 64., 23., 71., 38., 65., 55., 41., 59., 48., 71., 35., 56., 40., 58., 44., 80., 55., 37., 74., 51.,
58., 50., 60., 44., 57., 50., 45., 25., 59., 50., 71., 56., 74., 50., 58., 45., 54., 36., 54., 48., 55.,
45., 57., 50., 62., 44., 64., 43., 52., 38., 60., 55., 41., 53., 49., 34., 35., 54., 45., 68., 38., 50.,
60., 39., 59., 40., 57., 54., 23.
C>MARQT D70
NO.OF DATA---NOB=
OPTION FOR MEAN ESTIMATION---IMEAN=(YES=1, NO=0)

70 1

AR ORDER--N=
MA ORDER--M=
2 0

MAX. NO. OF ITERATIONS---ITMAX=
TOLERANCE FOR THE CHANGE OF PARAMETERS---TOL(1)=
CONVERGENT CRITERION---TOL(2)=
10 .001 .001
INPUT INITIAL VALUES OF PHI (S) AND THETA(S)
-.33847 . 19012
   -.33847        .19012
```

```
SUM  OF  SQUARE=        .79658140E+04    LAMBDA=     .450E+01
SUM  OF  SQUARE=        .79657960E+04    LAMBDA=     .250E+01
SUM  OF  SQUARE=        .79657960E+04    LAMBDA=     .139E+01
--------------------------------------------------------------

      ARMA(2, 0)
             53.3771
  MEAN=      51.2043         VAR=     140.637      AVG=     51.1429
             49.0314
                  AUTOREGRESSIVE  PARAMETERS
     -. 105833       .426506
     -. 345165       .187196
     -. 584497      -.052114

    SIGMA  E-SQUARE=    113.80
   MAGNITUDE  OF  GRADIENT  VECTOR
        - .00307        -.00163        .00104
   AIC=     337.40920            BIC=   347.15470
Stop-Program  terminated.
```

20. 直接法的谱估计（PMPSD）

20.1 原理简介

首先将含有N个数据的序列 $x(i)$ 构成NSECT个含有L 个数据的子序列，相邻两段子序列有一半数据重叠。例如$N=32$的序列可构成7个长度为$L=8$的子序列：$\{x_1,x_2,\cdots,x_8\}$，$\{x_5,x_6,\cdots,x_{12}\}$，$\{x_9,x_{10},\cdots,x_{16}\}$，…，$\{x_{25},x_{26},\cdots,x_{32}\}$。

将序列扣除均值后，对每段子序列加窗处理，可供选择的窗函数有以下六种：

矩形窗(RECT)

$$w(i)=\begin{cases}1 & 1\leqslant i\leqslant L\\ 0 & \text{其他}\end{cases}$$

汉明窗(HAMG)

$$w(i)=\begin{cases}0.54-0.46\cos\left(\dfrac{2\pi i}{L-1}\right) & 1\leqslant i\leqslant L\\ 0 & \text{其他}\end{cases}$$

通用汉明窗(GHAM)

$$w(i)=\begin{cases}a-(1-a)\cos\left(\dfrac{2\pi i}{L-1}\right) & 1\leqslant i\leqslant L\\ 0 & \text{其他}\end{cases}$$

海宁窗(HANN)

$$w(i)=\begin{cases}0.5-0.5\cos\left(\dfrac{2\pi i}{L-1}\right) & 1\leqslant i\leqslant L\\ 0 & \text{其他}\end{cases}$$

三角（巴特利）窗(TRIA)

$$w(i)=\begin{cases}2i/L & 1\leqslant i\leqslant L/2\\ 2[1-(i-1)/L] & L/2+1\leqslant i\leqslant L\end{cases}$$

帕森-2窗(PARZ)

$$w(i)=\begin{cases}1-4[(L/2-i)/L]^2 & 1\leqslant i\leqslant L/2\\ 1-4[(i-L/2-1)/L]^2 & L/2+1\leqslant i\leqslant L\end{cases}$$

加窗后的第l段子序列记作$x_l^w(i)\equiv x_l(i)w(i)$，将相邻两段构成一组复数序列$x_l^w(i)+jx_{l+1}^w(i)$，共有KMX=(NSECT+1)/2组。利用FFT算法对KMX组复序列求DFT—记作$Z_l(k)$。若M为FFT的计算点数，并以$X_l^w(k)$表$x_l^w(i)$的DFT，则因（见附录一）

$$Z_l(k)Z_l^*(k)+Z_l(M-k)Z_l^*(M-k)$$
$$=2[|X_l^w(k)|^2+|X_{l+1}^w(k)|^2]$$

故可按上式左边求得$l=1,\cdots,KMX$所有各组的和式

$2\sum\limits_{l=1}^{NSECT}|X_l^w(k)|^2$，功率谱的估计为

$$S_x(k)=\frac{1}{U\cdot NSECT}\sum_{l=1}^{NSECT}|X_l^w(k)|^2$$

其中

$$U=\sum_{i=1}^{L-1}|w(i)|^2$$

对$S_x(k)$求离散富氏反变换可得自相关函数(若原始数据未作方差归一化处理则为自协方差）。

20.2 主程序功能和参数说明

计算时间序列的功率谱（以$20\log_{10}S_x(k)$的分贝值给出）和自相关函数并打印图形。

由用户给定的有关计算参数有：序列总长度；FFT计算长度M($=2^n$，n为非零正整数，且$L\leqslant M\leqslant 1024$)；窗函数的类型和长度（即分段子序列长度）；数据的采样频率；通用汉明窗的参数α；是否打印功率谱图和是否打印自相关图。

〔输入量〕

N——序列总长度

XD——序列数值

M——*FFT*计算长度

L——数据窗长度（偶数，≤M）

IWIN——窗函数类型标志，1=RECT，2=HAMG，
3=GHAM，4=HANN，5=TRIA，6=PARZ

FS——采样频率（*Hz*）

ALPHA——IWIN=2时ALPHA=0.54；IWIN=3时由用户给定；IWIN=4时ALPHA=0.5。

IP——是否作功率谱图标志，*IP*=1时作图

IC——是否作相关图标志，*IC*=1时作图。

〔输出量〕

XA——开始存贮分段读取的时序数值，最后存放相关函数值。

WD——数据窗函数的值。

XREAL，XIMAG——开始存放相邻子序列构成复序列时的实部和虚部，最后存放相关函数值。

REXMN，IMXMN——分段数组实部和虚部的均值。

XFR——功率谱的频率值，范围是0到$FS/2$。

SXX——最后存放功率谱的值。

ILAG——相关函数的迟后值。

20.3 子程序说明

子程序GETD，FFT见程序CORELA说明。子程序GRAPH见程序ARMASP说明。

20.4 程序运行及举例

```
C>PMPSD AR4
FFT LENGTH---M=
128
WINDOW TYPE---1=RECTANGULAR, 2=HAMMING
3=GENERALIZED HAMMING, 4=HANNING
5=TRIANGULAR, 6=PARZEN
6
WINDOW LENGTH---L=
64
MAX.NO.OF DATA---N=
1024
SAMPLING FREQ.IN HZ---FS=
1
WINDOW TYPE=PARZ
M= 128 NP=1024 L= 64 SAMPLING FREQUENCY= 1.0000
XMEAN= -.82782E-01
PLOT PSD (1=YES) ?
0
LOG POWER SPECTRUM
```

FREQ	DB	FREQ	DB
.000	18.2167	.008	18.4576
.016	18.0057	.023	17.2197
.031	22.0856	.039	25.3405
.047	26.1635	.055	30.2330
.063	34.1546	.070	35.2335
.078	42.8579	.086	50.8698
.094	63.5032	.102	73.9340
.109	78.5850	.117	77.8925
.125	73.8429	.133	73.5964
.141	65.0112	.148	70.7807
.156	58.3281	.164	42.8165
.172	36.9479	.180	24.6191
.188	24.4257	.195	19.9408

FREQ	DB	FREQ	DB
.203	11.5417	.211	7.9575
.219	8.3645	.227	1.1934
.234	−.1150	.242	−.4668
.250	−3.1843	.258	−5.7762
.266	−10.7718	.273	−13.3706
.281	−12.5309	.289	−15.2753
.297	−15.5223	.305	−18.3152
.313	−17.7355	.320	−18.0055
.328	−20.7701	.336	−23.8235
.344	−24.7158	.352	−21.7620
.359	−25.9396	.367	−28.1167
.375	−27.2198	.383	−26.7179
.391	−31.2004	.398	−27.4510
.406	−32.9510	.414	−31.9580
.422	−33.5293	.430	−27.3886
.438	−31.3172	.445	−31.4542
.453	−35.1044	.461	−28.9709
.469	−36.6187	.477	−32.7685
.484	−34.2898	.492	−30.8110
.500	−37.8698		

PLOT CORRELATION FUNCTION ? (1 FOR YES)

0

CORRELATION FUNCTION

LAG	CORR	LAG	CORR	LAG	CORR
0	.678E+03	1	.487E+03	2	.284E+02
3	−.424E+03	4	−.614E+03	5	−.456E+03
6	−.718E+02	7	−.299E+03	8	.458E+03
9	.354E+03	10	.941E+02	11	−.153E+03
12	−.264E+03	13	−.222E+03	14	−.966E+02
15	.216E+02	16	.836E+02	17	.945E+02
18	.845E+02	19	.714E+02	20	.481E+02

LAG	CORR	LAG	CORR	LAG	CORR
21	.606E+00	22	—.658E+02	23	—.118E+03
24	—.117E+03	25	—.524E+02	26	—.472E+02
27	.124E+03	28	.132E+03	29	.659E+02
30	—.325E+02	31	—.104E+03	32	—.110E+03
33	—.544E+02	34	.225E+02	35	.733E+02
36	.736E+02	37	.334E+02	38	—.154E+02
39	—.426E+02	40	—.381E+02	41	—.138E+02
42	—.985E+01	43	.188E+02	44	.127E+02
45	.125E+01	46	—.585E+01	47	—.531E+01
48	—.729E+00	49	.314E+01	50	.300E+01
51	.210E+00	52	—.230E+01	53	—.263E+01
54	—.110E+01	55	.667E+00	56	.140E+01
57	.935E+00	58	.195E+00	59	—.202E+00
60	—.311E+00	61	—.112E+00	62	.173E—01
63	.240E—01	64	.000E+00		

Stop—Program terninated

21. 相关函数加窗谱估计(CMPSD)

21.1 原理简介

对给定的自相关函数加窗处理，然后取其离散富氏变换得功率谱S_x，其分贝数为$20\log_{10}S_x$，本程序有八种窗函数可任选:矩形窗(RECT),汉明窗 (HAMG) ,通用汉明窗 (GHAM), 海宁窗(HANN),三角(巴特利)窗(TRIA),帕森—2窗(PARZ),修正凯泽窗 (KAIS) ，三系数窗(3--CO)。窗函数的表达式可见第四章表4.3。

本程序和CORELA连用构成间接法的谱估计。

21.2 主程序功能和参数说明

实现在指定迟后窗的功率谱估计的计算和作图。

〔输入量〕

N——相关函数点数

XA——开始存放输入的样本自相关值，最后存放输出的功率谱值。

FS——采样频率（HZ）

IWIN——窗函数类型标志，1=RECT，2=HAMG，3=GHAM，4=HANN，5=TRIA，6=PARZ，7=KAIS，8=3-CO

L——迟后窗宽度（即用于估计谱的相关函数点数）。

NFFT——FFT计算点数（NFFT/2+1等于功率谱的点数）

ALPHA——当IWIN=2时ALPHA=0.54；
当IWIN=3时ALPHA由用户给定；
当IWIN=4时ALPHA=0.5

BETA——当IWIN=2，3，4时BETA=1-ALPHA，
当IWIN=7时0≤BETA≤10，是由用户给定的修正凯泽窗参数
当IWIN=8时 0≤BETA≤0.045，是用户给定的三系数窗参数。

ID——是否打印功率谱图标志，ID=1打印，ID=0不打印。

〔输出量〕

XA——见上

XFR——功率谱的频率值，范围是0到$FS/2$

W——迟后窗数值

XC，XS，ZC，ZS为中间变量

21.3 子程序说明

子程序FFT和程序CORELA中相同。

子程序MODKAI用来生成$KAIS$窗函数，其中L为窗宽，BETA为窗函数的参数β，W为求得的窗函数值。本子程序要调用一个实函数过程$IO(X)$以便计算贝塞尔函数。

子程序GRAPH和程序ARMASP中相同。

21.4 程序运行及举例

对PMPSD程序运行举例中算得的自协方差函数（存放在文件U3中）用修正凯泽窗作功率谱估计。

```
C>CMPSD  U3
TOTAL  NO.  OF  CORRELATION———N=
64
SAMPLING  FREQUENCY  IN  Hz———FS=
1
WINDOW  TYPE————1=RECTANGULAR, 2=HAMMING,
3=GENERALIZED  HAMMING, 4=HANNING,5=TRIANGULAR,
6=PARZEN, 7=MODIFIED  KAISER, 8=3—COEFFICIENT
7
NO.  OF  CORRELATION  VALUES  USED———L=
64
FFT  SIZE———NFFT
128
WINDOW  TYPE=KAIS  NO.  OF  WINDOW  VALUES=  64
FFT  SIZE=  128
 SPECIFY  BETA  FOR  MODIFIED  KAISER  WINDOW
 (0<=BETA<=10)
5
PLOT  POWER  SPECTRUM?  (1=YES, 0=NO)
0
LOG  POWER  SPECTRUM
```

FREQ	DB	FREQ	DB
.000	18.3257	.008	18.3009
.016	17.9428	.023	18.8175
.031	22.1414	.039	24.9019
.047	27.1471	.055	30.6431
.063	33.6840	.070	37.5947
.078	44.7392	.086	55.2244
.094	66.2472	.102	73.9469

FREQ	DB	FREQ	DB
.109	77.5001	.117	77.2682
.125	74.9357	.133	74.0136
.141	73.8326	.148	70.5003
.156	62.2034	.164	49.2100
.172	37.6264	.180	29.4156
.188	23.6667	.195	20.0667
.203	13.8732	.211	9.0591
.219	7.0380	.227	3.2680
.234	.1622	.242	—.9408
.250	—2.4929	.258	—6.0012
.266	—9.1234	.273	—12.4893
.281	—13.2439	.289	—14.6180
.297	—16.0160	.305	—17.4163
.313	—17.8381	.320	—18.5226
.328	—20.5853	.336	—23.1980
.344	—23.6358	.352	—23.3497
.359	—25.1984	.367	—27.5278
.375	—27.2321	.383	—27.5818
.391	—28.6558	.398	—29.2541
.406	—30.7349	.414	—32.6127
.422	—31.2336	.430	—29.4970
.438	—30.0480	.445	—32.1755
.453	—32.3664	.461	—31.7363
.469	—33.1020	.477	—34.1495
.484	—32.3131	.492	—32.9523
.500	—33.9974		

Stop —Program terminated

22.极大熵谱估计的伯格算法（MEBURG）

22.1 原理简介

见第五章5.3节式（5.3.1),(5.3.8～10)。

以最终预报误差作为定阶准则，计算规范化的最终预报误差

（除以数据序列的样本方差）的对数值，即

$$\log_{10}[FPE(m)/(\sum_{t=1}^{N} X_t^2/N)]$$

（其中m为递推阶数），以能使该值为最小的m作为最佳阶数。

FPE(m)的计算公式可任选：

（1）赤池公式

当不扣除均值时 $FPE(m)=\frac{N+m}{N-m}P_m$

当扣除均值时 $FPE(m)=\frac{N+m+1}{N-m-1}P_m$

（2）修正公式

当不扣除均值时 $FPE(m)=\frac{N+2m}{N-2m}P_m$

当扣除均值时 $FPE(m)=\frac{N+2(m+1)}{N-2(m+1)}P_m$

式中P_m为残差方差。

22.2 主程序功能和参数说明

在规定最高阶数MLAST后，程序从1阶开始按伯格递推算法求出各种阶数下的自回归系数和残差方差，并根据指定的FPE算式指出在上述阶数范围内的最佳阶数。

其他任选项有：是否在数据序列中扣除均值；是否计算并打印功率谱；程序运行是否要重复。当需要计算功率谱时，尚需给定最低频率、频率间隔和计算点数。注意这里计算的是和阶数MLAST对应的模型的功率谱，故若要计算某一特定阶数下的功率谱，应取该阶数为最高拟合阶数。

〔输入量〕

N——序列数据总数

X——时序数据

IM——是否扣除均值标志，IM＝1扣除，IM＝0不扣除

MLAST——拟合计算的最高阶数。

IC——FPE算式标志，IC＝1用赤池公式，IC＝2用 修 正公式

IS——是否计算功率谱标志，IS＝1计算，IS＝0不计算。

FO——计算功率谱时的最小频率

DF——计算功率谱时的频率间隔

NF——计算功率谱的点数

〔输出量〕

Y——扣除均值后的时序数据

G——自回归系数$a_1, a_2, \cdots, a_m$(即$-\phi_1, -\phi_2, \cdots, -\phi_m$)

FPE——规范化最终预报误差的对数值

PS——功率谱值

A——中间变量

22.3 子程序说明

子程序BURG实现伯格方法的递推计算。

M——时序数据个数

F——时序数据

LG——拟合最高阶数＋1

LEXT——LG＋1

PM——递推式(5.3.9)中的P_M

IC，IM，FPE意义同前。G和A为中间变量。

子程序OUTPUT将每次递推计算结果显示出来，内 容包括FPE的规范化对数值，P_M值和自回归系数$\phi_1, \cdots, \phi_m$。本 子 程序的参数有

NN——递推拟合阶数

G——自回归系数a_i

GG——自回归系数ϕ_i

PM和FPE意义同上

子程序GRAPH见程序ARMASP说明

22.4 程序运行及举例

对$D20$中的20个数据作极大熵谱估计。

```
C>TYPE  D20
   .0141       1.0609       1.7826       2.6804       3.0536
  2.9605       2.7524       2.1767       1.6413       1.0371
   .1217       -.9359      -1.8501      -2.5495      -2.5454
 -2.9358      -3.0443      -2.2961      -1.7226       -.9031
C>MEBURG  D20
NUMBER  OF  OBSERVATIONS  N=?  (<=1024)
20
SUBTRACTS  MEAN-----IM=?  (YES=1, NO=0)
1
HIGHEST  ORDER---MLAST=  (<40)
8
FORMULA  FOR  ORDER  SELECTION-----IC=1  OR  2
1
AR  ORDER=1     FPE=      -.984075600     P=      .421121600
PHI=         .951991600
AR  ORDER=2     FPE=      -1.650232000    P=      .081929210
PHI=        1.806374000
PHI=        -.897468600
AR  ORDER=3     FPE=      -1.691769000    P=      .067010710
PHI=        1.423406000
PHI=        -.126652100
PHI=        -.426720300
AR  ORDER=4     FPE=      -1.809260000    P=      .045883480
PHI=        1.183803000
PHI=        -.197767200
```

```
PHI=          .372521000
PHI=        -.561499200
AR  ORDER=5     FPE=     -1.762722000     P=        .045668790
PHI=         1.145395000
PHI=        -.172285700
PHI=          .358983200
PHI=        -.480523700
PHI=        -.068402880
AR  ORDER=6     FPE=     -1.800937000     P=        .030258540
PHI=         1.105660000
PHI=        -.451417900
PHI=          .567529400
PHI=        -.580603000
PHI=          .596947400
PHI=        -.580891700
AR  ORDER=7     FPE=     -1.838287000     P=        .030232160
PHI=         1.088510000
PHI=        -.433793900
PHI=          .550387900
PHI=        -.563847500
PHI=          .583619900
PHI=        -.548248700
PHI=        -.029523530
AR  ORDER=8     FPE=     -1.782487000     P=        .030210230
PHI=         1.089305000
PHI=        -.419075800
PHI=          .534667000
PHI=        -.548659300
PHI=          .568794200
PHI=        -.536563700
PHI=        -.058844570
PHI=          .026936850

  OPTIMUM  ORDER=    6

CALCULATES  PSD?  (YES=1, NO=0)
0
```

```
MORE  FIT? (1=YES, 0=NO) =
0
Stop —Program terminated.
```

23 预报的逆函数法 (INVERF)

23.1 方法简介

平稳序列的逆函数预报方法见第六章6.3节。

23.2 主程序功能和参数说明

对阶数及参数为已知的*ARMA*序列用逆函数法进行预报。

〔输入量〕

IP,IQ——模型的*AR*及*MA*阶数(n和m)

F,C——模型的*AR*及*MA*参数(φ_i和θ_i)

K,*KK*——预报的起始及终止时刻。

W——观察序列$\{x_t\}$。

EPS——预报精度。

L——最大预报步数。

〔输出量〕

WK——存放时刻*K*1 ($K \leq K1 \leq KK$) *L*步以内的预报值

$\hat{x}_{k1}(1)$, $\hat{x}_{k1}(2)$, $\cdots \hat{x}_{k1}(L)$。

23.3 程序运行及举例

对于参数为$\varphi_1=0.8$, $\theta_1=-0.4$的ARMA(1,1)模型，由计算机产生仿真数据，建立数据文件FORECA.0

```
 1.880577,  2.779158,  3.792327,  4.159847,  3.644195,
 4.102682,  4.034015,  2.218640,  1.264800,   .508673,
  .335562, -1.594151, -2.162143,  -.808656, -1.106485,
-1.559811, -1.456013,  -.019296,  1.203465,  1.614709,
 2.654008,  1.724404,  1.153256,   .299612,   .741279,
 1.722999,   .909865,  1.157582,   .344194,  -.636438,
 -.236183,  -.085021,   .601703,  1.619774,  3.202307,
 4.656597,  4.945848,  3.338663,  2.313548,  1.988774,
```

```
1.924376,  1.579144,  1.124817,  1.727150,  3.185822,
3.603831,  2.576014,  -.739002,  -1.407799,  -.728453,
.313450,  1.285756,  2.686826,  4.350112,  5.429370,
```

用逆函数法进行平稳预报的运行结果如下：

```
C>INVERF FORECA.0
 AR-AND MA-ORDER OF THE MODEL IP=?  IQ=?
1, 1
 AR-PARAMETERS OF THE MODEL   :
.8
 MA-PARAMETERS OF THE MODEL   :
-.4
 START POINT AND END POINT OF FORECAST K=?  KK=?
50, 55
 THE STEP OF FORECAST L=?
3
 PRECISION OF FORECAST EPS=?
.001
FORECAST        FORECAST STEP
  TIME        0        1        2        3
   50     -.7285   -.4596   -.3677   -.2942
   51      .3135    .5603    .4482    .3586
   52      1.286    1.319    1.055    .8440
   53      2.687    2.697    2.157    1.726
   54      4.350    4.142    3.313    2.651
   55      5.429    4.859    3.887    3.110
  Stop-Program terminated.
```

24. 预报矢量法（VECTF）

24.1 方法简介

用预报矢量法对平稳序列进行预报，具体方法见第六章6.3节。

24.2 主程序功能和参数说明

对阶数及参数为已知的*ARMA*序列，用预报矢量法计算某

时刻L步以内的预报值。

〔输入量〕

IP,IQ——模型的AR及MA阶数（n和m）

F,C——模型的AR及MA参数（φ_i和θ_i）

W——观察序列$\{x_t\}$

K,KK——进行矢量外推预报的起始及终止时刻。

KI——进行预报的时刻$K1\in〔K,KK〕$

WW——$K1$时刻前的L步观察值$x_{k1-1},\cdots,x_{k1-L}$,

WK——$K1-1$时刻的L步预报值$\hat{x}_{k1-1}(1),\cdots,\hat{x}_{k1-1}(L)$,

WW1——观察值x_{k1}

〔输出量〕

WW——观察值$x_{k1},x_{k1-1},\cdots,x_{k1-L+1}$

WK——$K1$时刻的L步预报值$\hat{x}_{k1}(1),\cdots,\hat{x}_{k1}(L)$

24.3 子程序说明

24.3.1 子程序PRED

功能：计算 $\hat{x}_k^*(J)=\sum_{i=1}^{IP}\varphi_i\hat{x}_k(J-1)$，$\hat{x}_k^*(J)$ 是矢量预报所做预报值的组成部分

〔输入量〕

K——预报时刻

J——预报步数，$1\leq J\leq L$

IP,IQ,F,C,WK,WW含义同主程序

〔输出量〕

WJ——$\hat{x}_k^*(J)$的值

24.3.2 子程序GREENF

见本附录9。

24.4 程序运行及举例

对于23中的ARMA(1,1)模型,除建立数据文件 FORECA.0外，还要建立数据文件VECTF.0

—.4596, —.3677, —.2942,

这是用逆函数法（见本附录23.INVERF）进行平稳预报得到的预报值$\hat{x}_{50}(1)$,$\hat{x}_{50}(2)$,$\hat{x}_{50}(3)$,以此作为本节预报的初值。

调用本程序的运行结果是

```
C>VECTF  FORECA.0  VECTF.0
 AR- AND  MA- ORDER  OF  THE  MODEL  IP=?   IQ=?
1, 1
 AR- PARAMETERS  OF  THE  MODEL        :
.8
 MA- PARAMETERS  OF  THE  MODEL        :
-.4
 START POINT AND END POINT OF FORECAST K=?   KK=?
51, 55
 THE  STEP  OF  FORECAST  L=?
3
FORECAST         FORECAST  STEP
   TIME          0           1           2           3
    51         .3135       .3761       .3009       .2407
    52         1.286       1.392       1.114       .8912
    53         2.667       2.667       2.134       1.707
    54         4.350       4.153       3.323       2.658
    55         5.429       4.854       3.883       3.107
 Stop—Program terminated.
```

25 多变量自回归模型的参数估计（MVAR）

25.1 方法简介

多变量自回归模型的建模和参数估计方法见第七章7.4节。

25.2 主程序功能和参数说明

本程序可对多变量观察序列采用FPE准则寻求使FPE值为最小的多变量自回归模型，并能计算出当只取部分变元时对应于

不同变元取法（规定为前若干个变元）所产生模型的FPE值（即相应的$FPE_{0,1,\cdots,p}$值）。

〔输入量〕

N——每种变元的观察数据个数。

K——观察的变元个数。

M——拟合AR模型时的最高阶数。

Z——多变量观察序列，即$\{X_t\}$

L——只考虑部分变元时的变元个数，这是指只考察第1至第L个变量观察序列（$L\leqslant K$）。

〔输出量〕

RO——多变量观察序列的方差

R——多变量观察序列的相关阵元素

FPEO——最佳模型的FPE值。

MO——考虑前I个变元($I=1,2,\cdots,L$)模型的最佳阶数。

HA——考虑前I个变元时，最佳模型的参数。

25.3 子程序说明

25.3.1 子程序MULCOV

功能：计算多变量观察序列的相关阵

〔输入量〕

N,K,M,Z含意见主程序说明

〔输出量〕

R0,R——含意同主程序

25.3.2 子程序ABP

功能：递推计算$A_{IP+1,IP+1}$和$B_{IP+1,IP+1}$（A、B分别对对应于式(7.4.11)中的φ和ψ下同）。

〔输入量〕

N,M,K,L,RO,R 含意同上

IP——逆推阶数，IP$\leqslant M$。

A,B——分别存放系数阵 $A_{IP,1}$,…, $A_{IP,IP}$和$B_{IP,1}$,…, $B_{IP,IP}$。

〔输出量〕

A,B——第$IP+1$个二维数组,分别存放$A_{IP+1,IP+1}$,和$B_{IP+1,IP+1}$。

FPE——第$IP+1$阶模型的FPE值。

ISW——出口标记，含意见18。

25.3.3 子程序ABJ

功能：计算$A_{IP+1,J}$,和$B_{IP+1,J}$,$1\leqslant J\leqslant IP$

〔输入量〕

J——1至IP之间的某正整数。

A,B——子程序ABP的输出量

M,L,IP——含意同上

〔输出量〕

A,B——第J个二维数组，分别存放$A_{IP+1,J}$，和$B_{IP+1,J}$。

25.3.4 子程序ABPJ

功能：由$A_{IP,J}$,$B_{IP,J}$($1\leqslant J\leqslant IP$) 递推计算$A_{IP+1,J}$, $B_{IP+1,J}$ ($1\leqslant J\leqslant IP+1$)

〔输入量〕

N,M,K,L,IP,R,RO,FPE含意同前。

A,B——分别存放$A_{IP,J}$和$B_{IP,J}$ ($1\leqslant J\leqslant IP$)

〔输出量〕

A,B——分别存放$A_{IP+1,J}$和$B_{IP+1,J}$ ($1\leqslant J\leqslant IP+1$)

25.3.5 子程序LDLT

基本上和18.3.2所说明的子程序相同，但本程序功能较强，X,B都是N×M矩阵。

25.4 程序运行及举例

用第七章表7.2中的多元气象资料建立数据文件MYAR.0,

使用本程序建立多变量自回归模型。

下列打印结果中，为节省篇幅，省略相关阵（可见表7.3）。MAX-DIM和DIM分别与式(7-4-20)中的p,l对应，AR-ORDER是使$FPE_{n,l,p}$为最小的阶数n，MUL. AR-PARAMETERS是相应的多变量AR模型的系数矩阵，每个阵均按列打印，并依照$\varphi_0,\varphi_{n,1},\varphi_{n,2},\cdots\varphi_{n,n}$顺序排列（对照7.4.4节举例）。R.M.S和MEAN分别为各变量样本的均方根值和均值。

```
C>MVAR  MVAR.0
  N=?   K=?   M=?
  120  4  20

    R.M.S.        .   64.864   7.958   8.655   8.869
    MEAN          .   84.545  16.031  15.640  16.392
MAX-DIM   DIM    AR-      FPE      MUL.AR-
                 ORDER             PARAMETERS

   4          1    1   .361E+04
                          78.560
                            .071

   4          2    3   .237E+05
                          74.668  -.144
                            .033    .002   2.109   .413
                          -.105   -.009   2.260   .469
                            .025  -.004  -3.502   .187

   4          3    5   .420E+05
                         105.759  -3.042  26.014
                            .003    .005   -.002   2.651   .279  -.199   4.995  -.182   .466
```

				−.101	−.005	.002	3.626	.316	−.156	2.289	.043	−.058
				.034	−.003	−.001	−1.808	.108	−.139	−2.303	.087	−.325
				.034	−.002	−.001	−3.991	−.072	−.095	−3.824	.088	.137
				−.007	.008	−.003	−3.617	−.019	.133	.902	.539	−.387
4	4	3	.607E+05									
				79.654	−11.968	38.484	38.852					
				.033	.002	.000	−.001	2.109	.413	−.414	−.433	6.192
				−.377	.500	.215	−.955	.032	.004	.317		
				−.105	−.009	.005	.005	2.260	.469	−.423	−.511	1.280
				−.210	−.069	196	1.151	.690	−.374	−.764		
				.025	−.004	.001	.005	−3.502	.187	−.254	−.141	−4.876
				.249	−.477	.041	−2.977	.322	.056	−.337		
3	1	3	.347E+04									
				89.519								
				.025								
				−.095								
				.012								
3	2	6	.251E+05									

				274.859	11.038							
				−.007	.005	2.107	.381					
				−.102	−.003	.097	.007					
				.028	.001	−3.459	−.265					
				.017	−.002	−3.319	−.127					
				−.018	.005	−1.299	.087					
				−.148	.001	−4.787	.194					
3	3	5	.546E+05									
				244.147	1.952	25.407						
				.005	.006	−.002	.712	.543	−.338	3.788	−.166	.473
				−.095	−.005	.001	1.197	−.020	−.105	.931	.048	−.160
				.020	.000	−.001	−1.726	−.232	−.057	−5.003	.128	−.163
				.020	−.003	−.001	−2.350	−.080	−.044	−4.297	−.156	.062
				−.026	.006	−.003	−2.436	.220	.055	−.494	.582	−.301
2	1	2	.351E+04									
				86.061								
				.059								
				−.077								

```
2          2    6   .269E+05
                         107.950  15.417
                            .037    .003  -2.631     .684
                          -.103   -.007   -.046    -.003
                          -.015    .000   1.385    -.341
                          -.030   -.001    .121     .058
                          -.040    .007    .431    -.111
                          -.113    .001    .673    -.269

1          1    7   .384E+04
                         130.137
                            .130
                          -.060
                          -.003
                          -.076
                          -.141
                          -.237
                          -.152
```

Stop-Program terminated.

26 混合回归模型及疏系数模型的估计（SMLR）

26.1 方法简介

混合回归和疏系数模型的变元挑选及建模方法 见 第八章8.2节。

26.2 主程序功能及参数说明

对多变量或单变量观察序列建立使BIC值为近似 最小的疏系数混合回归模型

〔输入量〕

N——每种变元的观察数据个数。

IS——观察值的变元个数。

X——多变量观察序列$\{X_t\}$。

M——对每一种变元的观察值所考虑的最大迟后步数。

KK——因变元观察值变量的标号。

〔输出量〕

II——最佳疏系数混合回归模型的阶数（参数个数）。

L——表明最佳模型对变元取用顺序的数组。

F——最佳疏系数混合回归模型的参数。

BICO——最佳模型的BIC值。

26.3 子程序说明

26.3.1 子程序SUBLSM

功能：按照混合回归模型构造相应的a_{ij}和$z_t$$(1\leqslant j\leqslant M*IS)$

26.3.2 子程序RECOV

功能：构造与正规方程相应的加边矩阵。

〔输入量〕

N1,N2——取用样本的起始和终止时刻。

MRE——模型的参数个数。

X,IS,N,M,KK含意同前。

〔输出量〕

S——与正规方程相应的MRE+1阶加边矩阵。

XX,ES——标准化线性方程组用到的中间变量。

26.3.3 子程序SELECT

功能：用消去变换法寻求使BIC值为近似最小的疏系数混合回归模型。

〔输入量〕

XX,ES,S含意见26,3,2。

〔输出量〕

II,FF,BICO——最佳疏系数混合回归模型的阶数、参数及BIC值。

L——最佳模型变元的取用顺序。

ISW——出口标记。(见18)

26.3.4 子程序ELIMIN

功能：施行(K,K)消去变换

〔输入量〕

S,MM——进行消去变换的矩阵及维数。

M——在S中施行消去变换的左上角子矩阵维数，$M \leqslant MM$

K——欲施行消去变换的对角线元素的标号。

〔输出量〕

S——对左上角$K \times K$子矩阵施行(K,K)消去变换后的矩阵。

26.4 程序运行及举例

用第八章表8-8中化肥年增产量数据建立数据文件SMLR.0，对该组数据建立疏系数自回归模型。调用本程序的运行结果列出如下：

```
C>SMLR SMLR.0
 NUMBER OF DATA IN EACH DIMENSION N=?
32
 DIMENSION OF OBSERVATIONAL DATA IS=? (IS<=5)
1
  MAXIMUM LAG M=?
8
 DIMENSIONAL INDEX OF OBSERVATIONS TAKEN TO BE
  DEPENDENT VARIABLE KK=?
1
 K   BIC      R.M.S.SELECTED ORDER OF VARIABLES PARAMETERS OF MODEL
 8 .9550E+01 .4108E+02
                     1 2 3 4 5 6 7 8
                              6.607 -.016 -.127 -.442 -.180 -.098 1.327 .656 1.342
 7 .9285E+01 .4109E+02
                     8 2 3 4 5 6 7
                              6.431 1.320 -.131 -.442 -.173 -.095 1.325 .630
 6 .9025E+01 .4118E+02
                     8 2 3 4 7 6
                              4.695 1.307 -.130 -.451 .157 .651 1.323
 5 .8782E+01 .4163E+02
                     8 2 3 6 7
                              2.978 1,308 -.164 -.474 1.251 .717
 4 .8554E+01 .4240E+02
```

```
                              8 7 3 6
                                              1.719  1.167   .656  -.547  1.273
3  .8429E+01  .4549E+02
                              8 6 3
                                              9.975  1.454  1.316  -.651
2  .8598E+01  .5660E+02
                              3 6
                                              30.469 -.429  1.161
1  .8527E+01  .6226E+02
                              6
                                              16.784 1.052
0  .8506E+01  .7030E+02
                                              37.075
```

OPTIMAL MODEL IS: K=3 BIC=.8429E+01 R.M.S= .549E+02

SELECTED ORDER VARIABLES 6 3 8

PARAMETERS OF MODEL 9.975 1.316 -.651 1.454

THE OPTIMAL S.M.L.R. MODEL IS:

```
X(K. 1) =9.975
          +(1.316)*X(K-6, 1)
          +(-.651)*X(K-3, 1)
          +(1.454)*X(K-8, 1)
```

Stop—Program terminated.

27 门限回归模型的估计(TREG)

27.1 方法简介

在门限分点给定时，门限回归（自回归）模型的建立及参数估计，参阅第八章8.3节。

27.2 主程序功能和参数说明

本程序可对给定的一组门限分点建立相应区间上的门限回归（自回归）模型。

〔输入量〕

N——每种变元观察值的数据个数。

IS——观察值的变元个数。

X——多变量观察序列，即$\{X_t\}$。

M——对门限自回归模型所取的最大迟后步数。

ID——门限元迟后量。

KK1——因变元观察值的变量标号。

KKD——门限元观察值的变量标号。

SP1，SP2——门限分点(SP1<SP2)，一般取0.1，0.2，…，0.9。

〔输出量〕

AIC——〔SP1,SP2〕内门限回归模型的AIC值。

MO——对于门限回归模型，其阶数MO=IS－1；对于门限自回归模型，MO是使AIC值为最小的自回归模型阶数(MO≤M)。

A，AO——〔SP1，SP2〕内门限回归（自回归）模型的参数，AO是常数项。

27.3 子程序说明

27.3.1 子程序ORDER

功能：找出把某一变元的观察数据按从小到大排列时的标号序列。

〔输入量〕

KKO——进行排序的观察序列的变元标号。

N,IS,X含意同前。

〔输出量〕

L——使第KKO个观察序列从小到大排列的标号序列。

27.3.2　子程序LSME

见18说明

27.3.3　子程序SUBLSM

功能：按照门限回归（自回归）模型构造a_{tj}和z_t。对于门限模型，$1\leqslant j\leqslant IS-1$；对于门限自回归模型，$1\leqslant j\leqslant M$。

27.4　程序运行及举例

用第八章表8.12给出的流量—雨量数据建立数据文件TREG.0，以流量为因变元，雨量为自变元建立门限回归模型，运行结果如下

```
C> TREG TREG.O
   NUMBER OF DATA IN EACH DIMENSION N=?
591
   DIMENSION OF OBSERVATIONAL DATA IS=?
2
   MAXIMUM LAG OF THRESHOLD VARIABLE ID=?
0
   THE DIMENSIONAL INDEX OF OBSERVATIONS TAKEN
           TO BE DEPENDENT VARIABLE KK1=?
1
   THE DIMENSIONAL INDEX OF
THRESHOLD VARIABLE KKD=?
1
   SWITCHING POINTS SP1=?  SP2=?
.2 .5
SWITCHING SWITCHING BIC S.D.  T.R.-   T.R.-
POINTS     POINTS               ORDER PARAME.
SP1 SP2    IN DATA
 .2  .5    315. 638.    9.1 92.1    1    430.58 2.26
Stop - Program terminated.
```

28. 双线性模型的估计（BILIM）

28.1 方法简介

用反复残差法建立双线性模型的具体方法见第八章8.4节。

28.2 程序功能和参数说明

当阶数给定时，本程序可用反复残差法求满足一定精度的双线性模型参数估计值。

〔输入量〕

N——每一变元观察数据的个数。

IS——观察数据的变元个数，这里IS=1。

X——观察序列 $\{x_t\}$，存放在X（1,J），J=1，…，N中。

IP，IQ——双线性模型的自回归及滑动平均参数。

L,KK——双线性模型的乘积交叉项阶数。

EPS——计算精度。

〔输出量〕

F,C——双线性模型满足给定精度要求的自回归、滑动平均参数。

CC——双线性模型的乘积交叉项参数。

28.3 子程序说明

28.3.1 子程序SUBLSM

功能：按照双线性模型构造a_{tj}，z_t $1\leqslant j\leqslant$IP+IQ+L＊KK

28.3.2 子程序LSME

见18

28.3.3 子程序RESID

功能：对已得出参数估值的双线性模型计算残差序列。

〔输入量〕

F,C,CC——已得到的双线性模型参数估计值。

N,IS,X,IP,IQ,L含义同前。

〔输出量〕

X (2,J) ,——1≤J≤N存放相应双线性模型的残差序列。

28.4 程序运行及举例

按照模型

$$x_t=0.4x_{t-1}+0.4x_{t-1}e_{t-1}+\varepsilon_t$$

产生随机序列 $\{x_t, 1\leqslant t\leqslant 400\}$，以此建立数据文件*BILIM*.0。为节省篇幅，仿真数据不予罗列。对该组数据用反复残差法建立双线性模型，调用本程序的运行结果是：

```
C>BILIM BILIM.0
 NUMBER  OF  DATA  N=?
400
AR—AND MA— ORDER      IP=?   IQ=?
1,0
   L=?     KK=?
```

29. 指数模型的估计（EXPOM）

29.1 方法简介

用近似的线性化方法求指数模型参数的估计值，见第八章8.4节。

29.2 程序功能和参数说明

本程序对任意固定γ（指数模型的指数系数），用线性最小二乘法计算相应的α_i，β_i的估计值。

〔输入量〕

N—观察值序列的数据个数

IS—这里IS=1

X—观察序列$\{x_t\}$

GAMA—给定的指数系数γ

IN—指数模型的阶数。

〔输出量〕

AIC—所得指数模型的AIC值。

F,FO—所得指数模型的参数α_1，…，α_{IN}，β_1 …，β_{IN}。

FO是常数项。

29.3 子程序说明

29.3.1 子程序*SUBLSM*

功能：按照指数模型构造a_{ij}，z_i，$1\leqslant j\leqslant z*IN$。

29.3.2 子程序LSME

见本附录18

29.4 程序运行及举例

按照模型

$$x_t=(0.6-0.2\exp(-3.0x_{t-1}^2))\ x_{t-1}+e_t$$

产生仿真数据序列$\{x_t,\ 1\leqslant t\leqslant 400\}$，并以此建立数据文件*EXPO.MO*.为节省篇幅，仿真数据不予罗列。对该组数据用线性化方法建立指数模型，调用本程序的运行结果是

```
C> EXPOM EXPOM.0
 NUMBER OF DATA  N=?
400
 IN=?   GAMA=?
1,3.,
 GAMA=     3.00 S.D.= .105E+01
 ALPH  :     .01      .65
 BETA  :              -.13
Stop -- Program terminated.
```

30 自适应AR模型的估计（ADAPAR）

30.1 方法简介

建立自适应*AR*模型的方法见第八章8.5节。

30.2 主程序功能和参数说明

本程序用自适应方法对观察数据建立*AR*模型，可计算出*L*步

以内AR模型参数及观察序列的预测值。

〔输入量〕

N—模型的AR阶数。

NN—观察数据的总长度。

L—预报最大步数

MEW—递推计算中选取的因子μ

K—当前时刻

XX—观察数据序列 $\{x_t\}$

X—K时刻观察数据的邻近历史值：x_K，x_{K-1}，…，x_{K-NL}，

$NL = N + L$

〔输出量〕

XL—K时刻观察序列预报值$\hat{x}_K(1)$，…，$\hat{x}_K(L)$

F—K时刻AR模型参数的预报值：φ_{K+1}，…，φ_{K+L} 。

运行中，递推AR参数时的初始φ值可以全部取为零（KK1＝0），或取为其他给定值（KK1＝1），这时接着要键入给定的初始参数。对x_t的初值也有两种选择，当$KK2=0$时全取为零；$KK2=1$时程序自动取头$N+L$个观察值作为初值，这时输出结果从$N+L+1$开始

30.3 程序运行及举例

按照二阶自回归模型

$$x_t - 0.5x_{t-1} - 0.1x_{t-2} = \varepsilon_t$$

利用$DAGENT$程序产生x_t序列（ε_t是均值为零，方差为0.2的正态白噪声），建立数据文件AR0501，下面是该序列的头64个数据

```
C>TYPE AR0501
 -.0340   .0513  -.2175   .2040   .2953  -.1845  -.2072  -.2459
 -.0978  -.3769  -.3084  -.1337  -.4350  -.0265  -.1967   .3200
  .1955  -.0401  -.1097   .0612  -.0057   .0258   .1132   .2947
```

.3263	.2376	.1683	−.0822	.0939	.1395	.0086	.1404
.0909	.3317	.2004	.0933	.4227	.1169	.2763	.3311
.2349	.0401	−.2098	.1739	.1336	−.0814	−.4274	.1023
.1868	.1590	.0337	−.0491	−.0044	.1183	−.0743	−.0739
.2674	.5337	−.0032	.0906	.0497	−.4766	−.5159	−.0914

运行结果如下

C>ADAPAR AR0501

NUMBER OF DATA---NN=?

430

AUTOREGRESSIVE ORDER---N=?

2

MAX-STEP OF FORECAST---L=?

2

PARAMETER---MEW=?

0.05

OPTION FOR INITIAL PARAMETERS

KK1=? (0=ALL ZERO, 1=NOT ZERO)

0

OPTION FOR STARTING POINT

KK2=? (0=FROM NO.1, 1=FROM NO.NL+1)

1

INITIAL VALUES OF X (I), I=1,…, 4:

−3.400000E−002

5.130000E−002

−2.175000E−001

2.040000E−001

5	−.0064	.0015	.2953	.0007	−.0022
6	−.0118	.0055	−.1845	− 0010	.0038
7	−.0079	−.0008	−.2072	0015	.0018
8	−.0028	.0038	−.2459	.0022	−.0001
⋮					
101	.1483	.0988	.1743	.0419	.0422
102	.1459	.0965	−.0972	.0036	.0026

103	.1438	.1003	.2242	.0241	.0225
104	.1439	.1003	.0286	.0259	.0266
105	.1427	.0910	—.3889	—.0493	—.0529
106	.1388	.0913	.0480	—.0323	—.0288
107	.1399	.0824	.1093	.0309	.0318
108	.1406	.0826	.0707	.0273	.0264
⋮					
271	.3213	.0936	—.2431	—.0881	—.0889
272	.3287	.0971	—.3922	—.1495	—.1525
273	.3306	.0983	—.2024	—.1028	—.1055
274	.3267	.0906	.0894	.0101	.0109
275	.3247	.0951	—.2119	—.0614	—.0603
276	.3364	.0932	—.6101	—.2178	—.2243
⋮					
342	.4504	.1270	—.0637	—.0101	—.0101
343	.4497	.1285	.0978	.0361	.0358
344	.4507	.1279	.1369	.0742	.0742
345	.4512	.1282	.1070	.0657	.0658
346	.4512	.1282	.1656	.0432	.0433
347	.4504	.1270	—.0711	—.0237	—.0237
348	.4526	.1250	—.3286	—.1563	—.1576
349	.4568	.1259	—.2864	—.1707	—.1722
⋮					
426	.4597	.0972	.3907	.2003	.2027
427	.4485	.0904	—.0856	—.0046	—.0031
428	.4472	.0962	.1460	.0586	.0571
429	.4427	.0988	—.2477	—.0998	—.0962
430	.4400	.1004	.0177	—.0160	—.0171

Stop—Program terminated.

本例中，因$KK2=1$，故x_t的4个初值为-0.0340，0.0513，-0.2175和0.2040。输出列表从$t=5$开始，在这一行中第一个和第二个数是φ_1和φ_2的递推值，第三个数是$t=5+1=6$时x_t的实际值，第四、五个数则分别为一步和二步预报值。其他各行的输出数据也按同样顺序排列。

1.1

TOLERANCE EPS=?

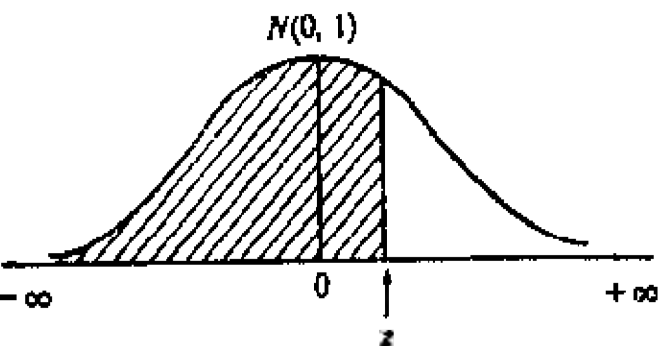
.001

ITERAT. No.	S.E.	F (I)	CC (I,J)
1	.127E+01	.249	.108
2	.118E+01	.156	.248
3	.114E+01	.237	.295
4	.112E+01	.296	.325
5	.111E+01	.334	.351
6	.109E+01	.364	.378
7	.106E+01	.393	.409
8	.104E+01	.429	.440
9	.105E+01	.479	.432
10	.105E+01	.501	.427
11	.105E+01	.509	.421
12	.105E+01	.508	.421
13	.105F+01	.507	.421
14	.105E+01	.506	.422

Stop-Program terminated.

附 录 三

表附Ⅲ－1 标准正态分布的累积分布函数表



本表给出 $\int_{-\infty}^{z} N(0,1)dz$ 的值

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
−3.5	0.00023	0.00022	0.00022	0.00021	0.00020	0.00019	0.00019	0.00018	0.00017	0.00017
−3.4	0.00034	0.00033	0.00031	0.00030	0.00029	0.00028	0.00027	0.00026	0.00025	0.00024

续表附Ⅲ—i

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
−3.3	0.00048	0.00047	0.00045	0.00043	0.00042	0.00040	0.00039	0.00038	0.00036	0.00035
−3.2	0.00069	0.00066	0.00064	0.00062	0.00060	0.00058	0.00056	0.00054	0.00052	0.00050
−3.1	0.00097	0.00094	0.00090	0.00087	0.00085	0.00082	0.00079	0.00076	0.00074	0.00071
−3.0	0.00135	0.00131	0.00126	0.00122	0.00118	0.000114	0.00111	0.00107	0.00104	0.00100
−2.9	0.0019	0.0018	0.0017	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
−2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
−2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
−1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
−1.7	0.0466	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
−1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0485	0.0495	0.0475	0.0465	0.0455

续表附Ⅲ—1

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
−1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
−1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0581
−1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
−1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1057	0.1038	0.1020	0.1003	0.0985
−1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
−1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
−0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
−0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
−0.7	0.2420	0.2389	0.2358	0.2327	0.2297	0.2266	0.2236	0.2207	0.2177	0.2148
−0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
−0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
−0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
−0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
−0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
−0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
−0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

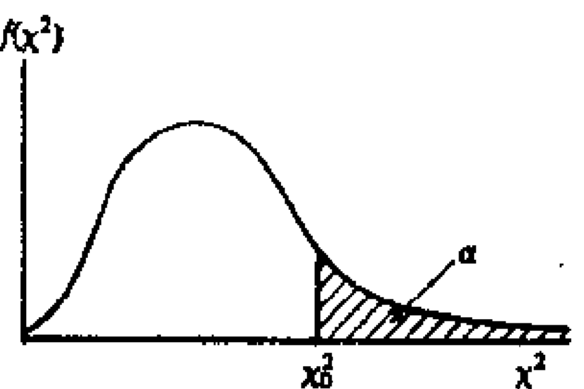
续表附Ⅲ—1

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
+0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
+0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
+0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
+0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
+0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
+0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
+0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
+0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
+0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8079	0.8106	0.8133
+0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
+1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
+1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
+1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
+1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
+1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
+1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
+1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
+1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
+1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706

续附表Ⅱ—1

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
+1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
+2.0	0.9773	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
+2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
+2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
+2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
+2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
+2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
+2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
+2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
+2.8	0.9974	0.9975	0.9976	0.9977	0.9077	0.9978	0.9979	0.9979	0.9980	0.9981
+2.9	0.9981	0.9982	0.9983	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
+3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.99900
+3.1	0.99903	0.99906	0.99910	0.99913	0.99915	0.99918	0.99921	0.99924	0.99926	0.99929
+3.2	0.99931	0.99934	0.99936	0.99938	0.99940	0.99942	0.99944	0.99946	0.99948	0.99950
+3.3	0.99952	0.99953	0.99955	0.99957	0.99958	0.99960	0.99961	0.99962	0.99964	0.99965
+3.4	0.99966	0.99967	0.99969	0.99970	0.99971	0.99972	0.99973	0.99974	0.99975	0.99976
+3.5	0.99977	0.99978	0.99978	0.99979	0.99980	0.99981	0.99981	0.99982	0.99983	0.99983

表附Ⅱ-2　χ^2分布表



本表给出χ_0^2值使 $P(\chi^2 \geq \chi_0^2)=\alpha$(阴影面积)

ν \ α	0.995	0.990	0.975	0.950	0.900	0.500	0.100	0.050	0.025	0.010	0.005
1	0.00	0.00	0.00	0.00	0.02	0.45	2.71	3.84	5.02	6.63	7.88
2	0.01	0.02	0.05	0.10	0.21	1.39	4.61	5.99	7.38	9.21	10.60
3	0.07	0.11	0.22	0.35	0.58	2.37	6.25	7.81	9.35	11.34	12.84
4	0.21	0.30	0.48	0.71	1.06	3.36	7.78	9.49	11.14	13.28	14.86
5	0.41	0.55	0.83	1.15	1.61	4.35	9.24	11.07	12.83	15.09	16.75

续表附Ⅲ—2

ν \ α	0.995	0.990	0.975	0.950	0.900	0.500	0.100	0.050	0.025	0.010	0.005
6	0.68	0.87	1.24	1.64	2.20	5.35	10.65	12.59	14.45	16.81	18.55
7	0.99	1.24	1.69	2.17	2.83	6.35	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	7.34	13.36	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	4.17	8.34	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	9.34	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	10.34	17.28	19.68	21.92	24.72	26.76
12	3.07	3.57	4.40	5.23	6.30	11.34	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	12.34	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	13.34	21.06	23.68	26.12	29.41	31.32
15	4.60	5.23	6.27	7.26	8.55	14.34	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	15.34	23.54	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	10.09	16.34	24.77	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	10.87	17.34	25.99	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	11.65	18.34	27.20	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	12.44	19.38	28.41	31.41	34.17	37.57	40.00

续表附Ⅲ—2

ν \ α	0.995	0.990	0.975	0.950	0.900	0.500	0.100	0.050	0.025	0.010	0.005
21	8.03	8.90	10.28	11.50	13.24	20.38	29.62	32.67	35.48	38.93	41.40
22	8.64	9.54	10.98	12.34	14.04	21.34	30.81	33.92	36.78	40.29	42.80
23	9.26	10.20	11.69	13.09	14.85	22.34	32.01	35.17	38.08	41.64	44.18
24	9.89	10.86	12.40	13.85	15.66	23.34	33.20	36.42	39.36	42.98	45.56
25	10.52	11.52	13.12	14.61	16.47	24.34	34.38	37.65	40.65	44.31	46.93
26	11.16	12.20	13.84	15.38	17.29	25.34	35.56	38.89	41.92	45.64	48.29
27	11.81	12.88	14.57	16.15	18.11	26.34	36.74	40.11	43.19	46.96	49.65
28	12.46	13.57	15.31	16.93	18.94	27.34	37.92	41.34	44.46	48.28	50.99
29	13.12	14.26	16.05	17.71	19.77	28.34	39.09	42.56	45.72	49.59	52.34
30	13.79	14.95	16.79	18.49	20.60	29.34	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	29.05	39.34	51.80	55.76	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	37.69	49.33	63.17	67.50	71.42	76.15	79.49
70	43.28	45.44	48.76	51.74	55.33	69.33	85.53	90.53	95.02	100.42	104.22
100	67.33	70.06	74.22	77.93	82.36	99.33	118.50	124.34	129.56	135.81	140.17

表附Ⅱ-3 F分布表



本表给出$F_{v_1v_2p}$使$P(F\geq F_{v_1v_2,0.95})=0.05$

v_2 \ v_1	1	2	3	4	5	6	7	8	9
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18

续表附Ⅲ—3

v_1 / v_2	1	2	3	4	5	6	7	8	9
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
12	4.75	2.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30

续表附Ⅲ—3

v_2 \ v_1	1	2	3	4	5	6	7	8	9
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88

v_2 \ v_1	10	12	15	20	24	30	40	60	120	∞
1	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50

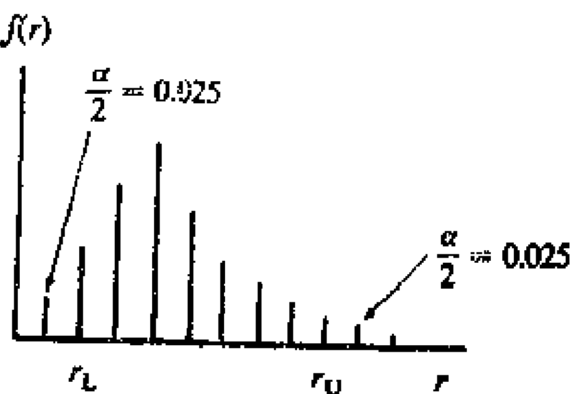
续表附Ⅲ—3

v_2 \ v_1	10	12	15	20	24	30	40	60	120	∞
3	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88

续表附 Ⅲ—3

ν_2 \ ν_1	10	12	15	20	24	30	40	60	120	∞
20	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

表附Ⅱ-4　游程检验用r分布



给定n_1和n_2，表中给出下限r_L和上限r_U使得它们左右对应的概率为$\alpha/2=0.025$ 即

$$P(r\leqslant r_L)+P(r\geqslant r_U)=0.05$$

n_1 \ n_2	L/U	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2	L											2	2	2	2
	U														
3	L					2	2	2	2	2	2	2	2	2	2
	U														
4	L				2	2	2	3	3	3	3	3	3	3	3
	U				9	9									

续表附Ⅲ—4

n_1 \ n_2	L/U	2	3	4	5	6	7	8	9	10	11	12	13	14	15
5	L			2	2	3	3	3	3	3*	4	4	4	4	4
	U			9	10	10	11	11							
6	L		2	2	3	3	3	3	4	4	4	4	5	5	5
	U			9	10	11	12	12	13	13	13	13			
7	L		2	2	3	3	3	4	4	5	5	5	5	5	6
	U				11	12	13	13	14	14	14	14	15	15	15
8	L		2	3	3	3	4	4	5	5	5	6	6	6	6
	U				11	12	13	14	14	15	15	16	16	16	16
9	L		2	3	3	4	4	5	5	5	6	6	6	7	7
	U					13	14	14	15	16	16	16	17	17	18
10	L		2	3	3	4	5	6	5	6	6	7	7	7	7
	U					13	14	15	16	16	17	17	18	18	18
11	L		2	3	4	4	5	5	6	[illegible]	7	7	7	8	8
	U					13	14	15	[illegible]	17	17	18	19	19	19

续表附Ⅱ—4

n_1 \ n_2	L / U	2	3	4	5	6	7	8	9	10	11	12	13	14	15
12	L	2	2	3	4	4	5	6	6	7	7	7	8	8	8
	U					13	14	16	16	17	18	19	19	20	20
13	L	2	2	3	4	5	5	6	6	7	7	8	8	9	9
	U						15	16	17	18	19	19	20	20	21
14	L	2	2	3	4	5	5	6	7	7	8	8	9	9	9
	U						15	16	17	18	19	20	20	21	22
15	L	2	3	3	4	5	6	6	7	7	8	8	9	9	10
	U						15	16	18	18	19	20	21	22	22

表附Ⅲ—5 调和分析中显著性检验Fisher检验表

(a) 周期分量的显著性参数 $\alpha=0.01$

n \ r	1	2	5	7	10
5	0.78853				
10	0.53584	0.30364			
15	0.40689	0.24190	0.11834		
20	0.32971	0.20223	0.10498		
25	0.27819	0.17448	0.09441	0.07353	
30	0.24124	0.15389	0.08591	0.06799	0.05156
35	0.2133	0.13796	0.07893	0.06324	0.04879
40	0.19156	0.12523	0.07309	0.05914	0.04625
45	0.17400	0.11482	0.06814	0.05558	0.04395
50	0.15954	0.10612	0.06387	0.05246	0.04186

(b) 周期分量的显著性参数　　$\alpha=0.01$

n \ r	1	2	5	10	25	50
100	0.08882	0.0618[illegible]	0.04024	0.02870	0.01659	0.00877
200	0.04853	0.0349[illegible]	0.02414	0.01826	0.01195	0.00786
300	0.03388	0.02485	0.01761	0.01367	0.00942	0.00663
400	0.02620	0.01943	0.01400	0.01104	0.00784	0.00573
500	0.02144	0.01603	0.01169	0.00932	0.00675	0.00505
600	0.01819	0.01368	0.01007	0.00810	0.00595	0.00453
700	0.01583	0.01196	0.00886	0.00717	0.00534	0.00411
800	0.01403	0.01065	0.00793	0.00646	0.00485	0.00377
900	0.01260	0.00960	0.00719	0.00588	0.00445	0.00349
1000	0.01145	0.00875	0.00658	0.00540	0.00411	0.00325
1100	0.01050	0.00805	0.00608	0.00500	0.00383	0.00305
1200	0.00970	0.00745	0.00565	0.00466	0.00359	0.00287
1300	0.00902	0.00694	0.00528	0.00437	0.00338	0.00271
1400	0.00843	0.00650	0.00495	0.00411	0.00319	0.00258
1500	0.00792	0.00612	0.00467	0.00389	0.00303	0.00245
1600	0.00746	0.00578	0.00442	0.00368	0.00288	0.00234
1700	0.00706	0.00547	0.00420	0.00350	0.00275	0.00224
1800	0.00670	0.00520	0.00400	0.00334	0.00263	0.0021[illegible]

续表附Ⅲ—5 (b)

n \ r	1	2	5	10	25	50
1900	0.00638	0.00496	0.00382	0.00320	0.00252	0.00206
2000	0.00609	0.00474	0.00365	0.00306	0.00242	0.00199
2100	0.00582	0.00453	0.00360	0.00294	0.00233	0.00192
2200	0.00568	0.00435	0.00337	0.00283	0.00224	0.00186
2300	0.00535	0.00418	0.00324	0'00273	0.00217	0.00179
2400	0.00515	0.00402	0.00312	0.00263	0.00210	0.00174
2500	0.00496	0.00388	0.00301	0.00254	0.00203	0.00168
2600	0.00478	0.00375	0.00291	0.00246	0.00197	0.00163
2700	0.00462	0.00362	0.00282	0.00238	0.00191	0.00159
2800	0.00447	0.00351	0.00273	0.00231	0.00185	0.00154
2900	0.00433	0.00340	0.00265	0.00225	0.00180	0.00150
3000	0.00419	0.00330	0.00257	0.00218	1.00175	0.00147

(c) 周期分量的显著性参数 $\alpha=0.05$

r / n	1	2	5	7	10
5	0.68377				
10	0.44495	0.26511			
15	0.33461	0.21016	0.10738		
20	0.27040	0.17547	0.09559	0.07324	
25	0.22805	0.15139	0.08612	0.06768	0.05008
30	0.19784	0.13360	0.07846	0.06274	0.04777
35	0.17513	0.11986	0.07215	0.05847	0.04540
40	0.15738	0.10890	0.06687	0.05475	0.04315
45	0.14310	0.09983	0.06238	0.05150	0.04108
50	0.13135	0.09244	0.05851	0.04865	0.03918

(d) 周期分量的显著性参数 $\alpha=0.05$

r / n	1	2	5	10	25	50
100	0.07378	0.05425	0.03704	0.02702	0.01584	0.00827
200	0.04074	0.03098	0.02234	0.01726	0.01160	0.00760
300	0.02861	0.02211	0.01635	0.01296	0.00909	0.00644
400	0.02222	0.01735	0.01303	0.01048	0.00767	0.00557
500	0.01825	0.01435	0.01090	0.00888	0.00652	0.00492

续表(d)

n \ f	1	2	5	10	25	50
600	0.01552	0.01228	0.00940	0.00770	0.00575	0.00441
700	0.01353	0.01075	0.00829	0.00683	0.00516	0.00401
800	0.01201	0.00958	0.00743	0.00615	0.00469	0.00368
900	0.01081	0.00865	0.00674	0.00561	0.00431	0.00341
1000	0.00984	0.00790	0.00617	0.00516	0.00398	0.00318
1100	0.00904	0.00727	0.00570	0.00478	0.00371	0.00298
1200	0.00836	0.00674	0.00530	0.00445	0.00348	0.00280
1300	0.00778	0.00628	0.00496	0.00417	0.00327	0.00265
1400	0.00728	0.00589	0.00466	0.00393	0.00310	0.00252
1500	0.00684	0.00554	0.00440	0.00372	0.00294	0.00240
1600	0.00645	0.00524	0.00416	0.00353	0.00279	0.00229
1700	0.00611	0.00497	0.00395	0.00336	0.00267	0.00219
1800	0.00580	0.00472	0.00377	0.00320	0.00255	0.00210
1900	0.00553	0.00450	0.00360	0.00306	0.00246	0.00202
2000	0.00527	0.00430	0.00344	0.00294	0.00235	0.00195
2100	0.00505	0.00412	0.00330	0.00282	0.00226	0.00188
2200	0.00484	0.00396	0.00318	0.00271	0.00218	0.00181
2300	0.00465	0.00380	0.00306	0.00261	0.00211	0.00175
2400	0.00447	0.00366	0.00295	0.00252	0.00204	0.00170

续表 (d)

r / n	1	2	5	10	25	50
2500	0.00431	0.00353	0.00285	0.00244	0.00197	0.00165
2600	0.00416	0.00341	0.00275	0.00236	0.00191	0.00160
2700	0.00402	0.00330	0.00267	0.00229	0.00186	0.00155
2800	0.00389	0.00320	0.00258	0.00222	0.00180	0.00151
2900	0.00377	0.00310	0.00251	0.00216	0.00175	0.00147
3000	0.00365	0.00301	0.00243	0.00210	0.00171	0.00144

附录四:程序文本

```
C                  *******************
C                  *                 *
C                  *   1.  DAGENT    *
C                  *                 *
C                  *******************


         DIMENSION X(2048)
         WRITE(*,10)
10       FORMAT(' AMPLITUDES OF TWO SINEWAVES---'
     *   'A1,A2=?'/)
         READ(*,*) A1,A2
         WRITE(*,20)
20       FORMAT(' FREQUENCIES OF TWO SINEWAVES---'
     *    'f1,f2(IN Hz)=?'/)
         READ(*,*) F1,F2
         WRITE(*,30)
30       FORMAT(' TYPE OF ADDED PROCESS---'
     *   '1=WHITE NOISE, 2=ARMA PROCESS'/)
         READ(*,*) IA
         WRITE(*,40)
40       FORMAT(' AVERAGE OF WHITE NOISE'
     *    '---AVG=?'/)
         READ(*,*) AVG
         WRITE(*,50)
50       FORMAT(' VARIANCE OF WHITE NOISE'
     *   '---VAR=?'/)
         READ(*,*) VAR
         WRITE(*,70)
70       FORMAT(' NO OF SIMULATE DATA---N=?'/)
         READ(*,*) NOB
         IF(IA.EQ.2) GOTO 120
         WRITE(*,60)
60       FORMAT(' pdf OF WHITE NOISE---ID=? ',
     *    '1=EXPONENTIAL',
     *   /'                                      2=UNIFORM',
     *   /'                                      3=NORMAL')
         READ(*,*) ID
         GOTO (2,3,1),ID
1        A=VAR
         B=AVG
         GOTO 100
2        WRITE(*,80)
```

```
80        FORMAT(' PARAMETER OF EXPONENTIAL pdf'
     *    '---LAMDA='/)
          READ(*,*) A
          GOTO 100
3         WRITE(*,90)
90        FORMAT(' LOWER AND UPPER LIMIT OF
     +    'UNIFORM pdf='/)
          READ(*,*) A,B
100       RN1=2.*(2.**14-2.)+1.
          G=.0
          PI=8.*ATAN(1.)
          DO 110 I=1,NOB
          XN=FLOAT(I)
          W=A1*COS(F1*XN*FI)+A2*COS(F2*XN*PI)
          CALL RANDOM(A,B,ID,G,RN1,SN)
          X(I)=W+SN
110       CONTINUE
          GOTO 150
120       CALL SML(NOB,VAR,AVG,X)
          DO 130 I=1,NOB
          XN=FLOAT(I)
          W=A1*COS(F1*XN*PI)+A2*COS(F2*XN*PI)
          X(I)=X(I)+W
130       CONTINUE
150       S=0.0
          DO 190 I=1,NOB
  190     S=S+X(I)
          XAVG=S/FLOAT(NOB)
          T=0.0
          DO 200 I=1,NOB
  200     T=T+(X(I)-XAVG)*(X(I)-XAVG)
          XVAR=T/FLOAT(NOB-1)
          WRITE(*,140)XAVG,XVAR
140       FORMAT(' XAVG=',F12.5,4X,'XVAR=',F12.5,/)
          OPEN(1,FILE='      ',STATUS='NEW')
          ND=NOB/8
          DO 160 I=1,ND
          I1=8*I-7
          I2=8*I
160       WRITE(1,170) (X(J),J=I1,I2)
          IF(MOD(NOB,8).EQ.0) GOTO 180
          ND1=ND+1
          NI=8*ND+1
```

```
      WRITE(1,170) (X(J),J=NI,NOB)
170   FORMAT(8F10.4)
180   STOP
      END
C*************************************************************
      SUBROUTINE BML(NOB,VAR,AVG,X)
      DIMENSION AT(2048),X(2048)
      DIMENSION PHI(20),THETA(20)
      WRITE(*,3)
3     FORMAT(' AR ORDER---n=?     MA ORDER---
     +'m=?'/)
      READ(*,*) NAR,NMA
      WRITE(*,4)
4     FORMAT(' AR PARAMETERS---PHI(1),....,'
     +,'PHI(n)=?'/)
      READ(*,*) (PHI(I),I=1,NAR)
      WRITE(*,5)
5     FORMAT(' MA PARAMETERS---THETA(1),....,'
     +,'THETA(m)=?'/)
      READ(*,*) (THETA(I),I=1,NMA)
      G=.0
      RN1=2.*(2.**14-2.)+1.
      DO 1 I=1,NOB
      CALL RANDOM(VAR,AVG,3,8,RN1,SN)
1     AT(I)=SN
      X(1)=AT(1)
      DO 45 N=2,NAR
      W=0.0
      K=N-1
      DO 15 I=1,K
      NI=N-I
15    W=W+PHI(I)*X(NI)
      W=W+AT(N)
      IF(NMA.EQ.0) GOTO 45
      IF(N-NMA) 20,20,25
20    L=N-1
      GOTO 40
25    L=NMA
40    DO 42 J=1,L
      NJ=N-J
42    W=W-THETA(J)*AT(NJ)
45    X(N)=W
      IP=NAR+1
```

```
      DO 65 N=IP,NOB
      W=0.0
      DO 50 I=1,NAR
      NI=N-I
   50 W=W+PHI(I)*X(NI)
      W=W+AT(N)
      IF(NMA.EQ.0) GOTO 65
      DO 60 J=1,NMA
      NJ=N-J
   60 W=W-THETA(J)*AT(NJ)
   65 X(N)=W
      RETURN
      END
C*********************************************
      SUBROUTINE RANDOM(A,B,ID,G,RN1,SN)
1     RN=RN1*(2.**8+3.)
      RN=RN-INT(RN/2.**15)*2.**15
      RSTN=RN
      RSTN=RSTN/2.**15
      IF(ID-2) 2,3,4
2     SN=-(1./A)*ALOG(RSTN)
      GOTO 5
3     SN=(B-A)*RSTN+A
      GOTO 5
4     IF(G-1.) 7,6,7
7     G=G+1.
      RN1=RN
      RSTN1=RSTN
      GOTO 1
6     Z=SQRT(-2.*ALOG(RSTN1))*A
      SN=Z*SIN(6.28318*RSTN)+B
      G=0.
5     RN1=RN
      RETURN
      END


C                  ******************
C                  *                *
C                  *   2.   SMSV    *
C                  *                *
C                  ******************
```

```
      REAL X(2048),E(2048)
      OPEN(1,FILE=' ')
      WRITE(*,101)
101   FORMAT(2X,'NUMBER OF DATA  N=?')
      READ(*,*)N
      READ(1,*)(X(I),I=1,N)
      CALL SMSV(N,X,E,EX,SIGM)
      WRITE(*,100)EX,SIGM
100   FORMAT(1X,'AVERAGE =',G12.4,
     * 3X,'VARIANCE =',G12.4)
      CLOSE(1)
      STOP
      END
C*********************************************************
      SUBROUTINE SMSV(N,X,E,EX,SIGM)
      REAL X(N),E(N)
      DO 10 I=1,N
10    E(I)=X(I)
      N1=N
30    N11=N1-1
      IF(N11.LE.0)GOTO  40
      DO 20 I=1,N11,2
      I1=I/2+1
20    E(I1)=0.5*(E(I)+E(I+1))
      N1=N1/2
      GOTO 30
40    EX=E(1)
      SIGM=0.
      DO 50 I=1,N
50    SIGM=SIGM+(X(I)-EX)**2
      SIGM=SIGM/FLOAT(N)
      RETURN
      END


C              ******************
C              *                *
C              *   3.  PDFH     *
C              *                *
C              ******************


      DIMENSION HIST(52),CURVE(52)
```

```
      DIMENSION X(2048),HIS(52)
      CHARACTER TEXT(24)
40    DO 10 I=1,52
      CURVE(I)=0.
10    HIS(I)=0.
      SX=0.
      SX2=0.
      HISO=0.
      HISE=0.
      WRITE(*,9999)
9999  FORMAT('** TITLE OF THIS SET OF DATA **')
      READ(*,9998) TEXT
9998  FORMAT(24A1)
      WRITE(*,9997)
9997  FORMAT(' NO. OF DATA---NDATA=?',/)
      READ(*,*) NDATA
      OPEN(1,FILE='    ')
      READ(1,*) (X(I),I=1,NDATA)
      CLOSE(1)
      DO 20 I=1,NDATA
20    SX=SX+X(I)
      XM=SX/NDATA
      DO 30 I=1,NDATA
      XA=X(I)-XM
30    SX2=SX2+XA*XA
      VAR=SX2/(NDATA-1)
      SIGMA=SQRT(VAR)
      WRITE(*,9995)
9995  FORMAT(' MIN. VALUE FOR ANALYSIS---'
     + ,'XBEG=?',/,' INTERVAL BETWEEN CLASSES',
     + '---DELX=?',/,' NUMBER OF CLASSES      ',
     + ' ---NX=?')
      READ(*,*) XBEG,DELX,NX
      XEND=XBEG+FLOAT(NX-2)*DELX
      DO 90 I=1,NDATA
      IF(X(I)-XBEG) 50,60,60
50    HISO=HISO+1.
      GO TO 90
60    IF(XEND-X(I)) 70,80,80
70    HISE=HISE+1.
      GO TO 90
80    J=IFIX((X(I)-XBEG)/DELX)+1
      HIS(J)=HIS(J)+1.
```

```
      K=J+1
      HIST(K)=HIS(J)
 90   CONTINUE
      HIST(1)=HISO
      HIST(NX)=HISE
      CALL GCURVE(XBEG,DELX,NX,XM,VAR,
     +  NDATA,CURVE)
      WRITE(*,9991) TEXT
9991  FORMAT(5X,24A1)
      WRITE(*,9994) NDATA,XBEG,XEND,DELX,XM,
     +  SIGMA
9994  FORMAT(' NDATA = ',I5/'  XBEG = ',F7.3,
     + '   XEND =',F7.3,'  DELTAX =',F7.3/
     + '  MEAN =',F7.3,'  STANDARD '
     + ,'DEVIATION =',F7.3/)
      CALL PRHIST(XBEG,DELX,NX,HIST,CURVE)
      PP=0.
      DO 100 I=1,NX
      P=((HIST(I)-CURVE(I))**2)/CURVE(I)
 100  PP=PP+P
      WRITE(*,9992) PP
9992  FORMAT(/,'  CHI SQUARE = ',G12.4)
      WRITE(*,9993)
9993  FORMAT(' MORE RUN ? (1 FOR YES)')
      READ(*,*) MR
      IF(MR.EQ.1) GOTO 40
      END
C*****************************************************
      SUBROUTINE GCURVE(XBEG,DELX,NX,XM,VAR,
     +  NDATA,CURVE)
      DIMENSION CURVE(1)
      SIGMA=SQRT(VAR)
      DO 10 J=1,NX
      X=XBEG+(J-1.5)*DELX
      C=EXP(-1.*(X-XM)**2/(2.*VAR))
      C=C/(2.5066*SIGMA)
 10   CURVE(J)=C*NDATA*DELX
      RETURN
      END
C*************************************************
      SUBROUTINE PRHIST(XBEG,DELX,NX,
     +      HIST,CURVE)
      DIMENSION HIST(1),CURVE(1)
```

```
      CHARACTER BLANK,CROSS,AST,ZLINE(50)
      DATA BLANK,CROSS,AST/' ','x','*'/
      SCFACT=1.
      HMAX=0.
      DO 2 J=1,NX
      IF(HIST(J)-HMAX) 2,2,1
1     HMAX=HIST(J)
2     CONTINUE
3     IF(HMAX-50.) 5,5,4
4     HMAX=HMAX/2.
      SCFACT=SCFACT/2.
      GOTO 3
5     DO 40 J=1,NX
      X=XBEG+(J-1.5)*DELX
      DO 10 I=1,50
10    ZLINE(I)=BLANK
      K=INT(HIST(J)*SCFACT+.5)
      IF(K) 25,25,15
15    DO 20 I=1,K
20    ZLINE(I)=CROSS
25    K=INT(CURVE(J)*SCFACT+.5)
      IF(K-50) 30,30,40
30    IF(K) 40,40,35
35    ZLINE(K)=AST
40    WRITE(*,1000) X,HIST(J),ZLINE
      RETURN
1000  FORMAT(1H ,F7.3,F5.1,2H  ,50A1)
      END


C              ******************
C              *                *
C              *   4.  TESTI    *
C              *                *
C              ******************


      CHARACTER*3 CHAR,IBLANK,LI(51)
      DIMENSION R(80),RR(80),VAR(80),S(80)
      DIMENSION X(1024)
      DATA CHAR/'*'/,IBLANK/' '/
      WRITE(*,1)
1     FORMAT('  NO. OF DATA --- N=?(<=1024)',/
```

```
     +  ' NO. OF AUTOCORRELATION VALUES USED---'/
     +  , ' KK=?')
        READ(*,*) N,KK
        OPEN(1,FILE='      ')
        READ(1,*) (X(I),I=1,N)
        CLOSE(1)
        XN=N
        CO=0.
        DO 103 I=1,N
103     CO=CO+X(I)*X(I)
C       CALCULATION OF R
        RKSUM=0.
        DO 104 K=1,KK
        CK=0.
        NN=N-K
        DO 105 J=1,NN
        JK=J+K
105     CK=CK+X(J)*X(JK)
        R(K)=CK/CO
104     RKSUM=RKSUM+R(K)*R(K)
        RKSUM=RKSUM*XN
        WRITE(*,111)RKSUM
111     FORMAT(5X,' Q=(SUM OF SQUARES OF AUTO',
     +    'CORRELATIONS)*N =',F12.5,/)
C       CALCULATION OF VAR AND VART
        VAR(1)=1./XN
        STSR=SQRT(VAR(1))
        S(1)=R(1)/STSR
        Z=2./XN
        DO 150 K=2,KK
        VAR(K)=VAR(K-1)+Z*(R(K-1)**2)
        STSR=SQRT(VAR(K))
        S(K)=R(K)/STSR
150     CONTINUE
C       PRINT AND PLOT THE AUTOCORRELATION
C            FUNCTION FROM 1 TO KK
        DO 160 I=1,KK
        IF(R(I).GT.0.5) GO TO 180
        RR(I)=R(I)
160     CONTINUE
        R1=-.5
        R2=.5
        GO TO 200
```

```
180       DO 190 I=1,KK
190       RR(I)=.5*R(I)
          R1=-1.
          R2=1.
200       WRITE(*,201)R1,R2
201       FORMAT(2X,'LAG K',3X,'AUTOCORRELATION',5
     +    X,'UNIFIED',10X,F4.1,22X,'0.0',21X,F4.1)
          WRITE(*,202)
202       FORMAT(47X,7H..I....,9(5HI....),3HI..)
          DO 205 I=1,51
205       LI(I)=IBLANK
          DO 280 I=1,KK
          IR=INT(RR(I)*50)
          IF(IR) 210,220,230
210       M=26-IABS(IR)
          K=26
          GOTO 240
220       M=26
          K=26
          GOTO 240
230       M=26
          K=IR+26
240       DO 250 J=M,K
250       LI(J)=CHAR
          WRITE(*,260)I,R(I),S(I),LI
260       FORMAT(3X,I2,8X,F8.4,8X,F8.4,12X,51A1)
          DO 270 J=M,K
270       LI(J)=IBLANK
280       CONTINUE
          STOP
          END

C                  ******************
C                  *                *
C                  *   5.  TESTP    *
C                  *                *
C                  ******************

          DIMENSION X(1030),A(515),B(515)
          DIMENSION C(515),F(515),G(515)
```

```
      WRITE(*,100)
      READ(*,*) N
100   FORMAT(' NO. OF DATA---N=?')
      OPEN(1,FILE='     ')
      READ(1,*) (X(I),I=1,N)
      CLOSE(1)
      SX=0.
      DO 10 I=1,N
10    SX=SX+X(I)
      XM=SX/FLOAT(N)
      AO=2.*XM
      WRITE(*,300) XM,AO
300   FORMAT(/,'  SAMPLE MEAN---XM=',G15.6,/,
     + 5X,' A(0)=',G15.6)
      DT=2.*3.1415967/FLOAT(N)
      SUM=.0
      IEO=MOD(N,2)
      NN=N-1
      IF(IEO.EQ.0) NN=NN-1
      M=INT(NN/2)
      DO 20 J=1,M
      DJ=DT*FLOAT(J)
      C1=COS(DJ)
      P=.0
      Q=.0
      C2=2.*C1
      DO 30 I=1,N
        K=N-I+1
        L=X(K)+C2*P-Q
        Q=P
        P=L
30    CONTINUE
      Q=2*(P*C1-Q)/FLOAT(N)
      P=2*P*SIN(DJ)/FLOAT(N)
      A(J)=Q
      B(J)=P
      F(J)=A(J)**2+B(J)**2
      C(J)=SQRT(F(J))
      SUM=SUM+F(J)
20    CONTINUE
      WRITE(*,400)
      DO 40 I=1,M
      G(I)=F(I)/SUM
```

```
      WRITE(*,500) I,A(I),B(I),C(I),F(I),G(I)
40    CONTINUE
      WRITE(*,600) SUM
      GM=.0
      M1=M-1
      DO 60 K=1,M1
      K1=K+1
       DO 50 J=K1,M
        IF(G(K).GT.G(J)) GO TO 50
        GM=G(K)
        G(K)=G(J)
        G(J)=GM
50    CONTINUE
60    CONTINUE
      WRITE(*,700) (I,G(I),I=1,M)
400   FORMAT(/,11X,'A(K)',8X,'B(K)',9X,'C(K)',
     + 9X,'C(K)**2',6X,'G')
500   FORMAT(1X,I4,1X,2F11.5,1X,3(F12.5))
600   FORMAT(/,'  SUM OF C(K)**2 =',F12.5)
700   FORMAT(/,1X,5(I4,F10.5))
      END

C                *****************
C                *               *
C                *   6.  POLYTR  *
C                *               *
C                *****************

      DOUBLE PRECISION C0,C1,C2,C3,C4,D0,D1,
     +  D2,D3,D4,F,G
      DIMENSION X(1024),F(5),G(5)
      OPEN(1,FILE='     ')
      WRITE(*,3)
3     FORMAT('  NO. OF DATA---N(<=1024)=?',/,
     +    '  ORDER OF POLYNOMIAL---K(<=4)=?')
      READ(*,*) N,K
      N=2*INT((N-1)/2)+1
      READ(1,*) (X(I),I=1,N)
      CLOSE(1)
      RN=N
```

```
      DO 10 I=1,5
10    F(I)=0.0
      DO 20 I=1,N
      A=I-1-(N-1)/2
      F(1)=F(1)+X(I)
      F(2)=F(2)+A*X(I)
      F(3)=F(3)+A**2*X(I)
      F(4)=F(4)+A**3*X(I)
      F(5)=F(5)+A**4*X(I)
20    CONTINUE
      G(1)=RN*(RN*RN-1.)/12.
      G(2)=G(1)*(3.*RN*RN-7.)/20.
      G(3)=G(1)*(3.*RN**4-18.*RN*RN+31.)/112.
      G(4)=G(1)*(5.*RN**6-55.*RN**4+239.*RN*RN
     +-381.)/960.
      L=K+1
      GOTO (5,15,25,35,45),L
5     CO=F(1)/RN
      GOTO 111
15    DO=F(1)/RN
      D1=F(2)/G(1)
      GOTO 99
25    DO=G(1)*F(3)-G(2)*F(1)
      DO=DO/(G(1)*G(1)-RN*G(2))
      D1=F(2)/G(1)
      D2=G(1)*F(1)-RN*F(3)
      D2=D2/(G(1)*G(1)-RN*G(2))
      GOTO 99
35    DO=F(1)*G(2)-G(1)*F(3)
      DO=DO/(RN*G(2)-G(1)*G(1))
      D1=G(2)*F(4)-G(3)*F(2)
      D1=D1/(G(2)*G(2)-G(1)*G(3))
      D2=G(1)*F(1)-RN*F(3)
      D2=D2/(G(1)*G(1)-RN*G(2))
      D3=G(2)*F(2)-G(1)*F(4)
      D3=D3/(G(2)*G(2)-G(1)*G(3))
      GOTO 99
45    D4=((RN*F(5)-G(2)*F(1))*(G(1)*G(1)-RN*
     1G(2))+(G(1)*F(1)-RN*F(3))*(G(2)*G(1)-RN*
     2G(3)))/((G(1)*G(2)-RN*G(3))**2-(G(2)*G(2)
     3-RN*G(4))*(G(1)*G(1)-RN*G(2)))
      D3=G(1)*F(4)-G(2)*F(2)
      D3=D3/(G(1)*G(3)-G(2)*G(2))
```

```
      D2=(G(1)*F(1)-RN*F(3)-D4*(G(1)*G(2)-RN*
    1 G(3)))/(G(1)*G(1)-RN*G(2))
      D1=G(3)*F(2)-G(2)*F(4)
      D1=D1/(G(1)*G(3)-G(2)*G(2))
      D0=(F(1)-D2*G(1)-D4*G(2))/RN
 99   B=-(N-1)/2
      C0=D0+B*D1+B**2*D2+B**3*D3+B**4*D4
      C1=D1+2.*B*D2+3.*B*B*D3+4.*B**3*D4
      C2=D2+3.*B*D3+6.*B*B*D4
      C3=D3+4.*B*D4
      C4=D4
111   CONTINUE
      WRITE(*,66) C0,C1,C2,C3,C4
 66   FORMAT('  THE POLYNOMIAL CONTENT IS:',/,
    1 '  (',G12.4,') + (',G12.4,')*t + (',G12.4,
    + ')*t**2 + (',G12.4,')*t**3 + (',G12.4,
    + ')*t**4')
      DO 18 I=1,N
      A=I-1
      X(I)=X(I)-C0-C1*A-C2*A**2-C3*A**3-C4*A**4
 18   CONTINUE
      WRITE(*,77) N
 77   FORMAT('   DATA AFTER TREND REMOVING :
    +  ,'( N =',I5')')
      WRITE(*,78) (X(I),I=1,N)
 78   FORMAT(5(1X,E14.7))
      STOP
      END


C               ******************
C               *                *
C               *   7.  INORD    *
C               *                *
C               ******************


      DIMENSION X(1024),XA(50)
      WRITE(*,10)
      READ(*,*) N,K
      OPEN(1,FILE='      ')
      READ(1,*) (X(I),I=1,N)
```

```
      CLOSE(1)
10    FORMAT(1X,'NO. OF DATA---N=?',/,
     *       1X,'NO. OF SEGMENTS---K=?')
      M=N/K
      DO 1 I=1,K
      SUM=0.0
      DO 2 J=1,M
      K1=(I-1)*M+J
 2    SUM=SUM+X(K1)
 1    XA(I)=SUM/M
      CALL INVT(K,XA,U)
      WRITE(*,20) U
20    FORMAT(4X,'U=',F8.4)
      WRITE(*,30)
30    FORMAT(4X,'(FOR A STATIONARY SERIES,',
     +  ' ABS(U) <= 2)')
      STOP
      END
C**************************************************************
      SUBROUTINE INVT(K,X,U)
      DIMENSION X(K)
      A=0
      K1=K-1
      DO 1 I=1,K1
      I1=I+1
      DO 2 J=I1,K
      IF(X(J).GT.X(I)) A=A+1.
2     CONTINUE
1     CONTINUE
      EA=K*K1/4.
      VA=((2*K+3)*K-5)*K/72.0
      U=(0.5+A-EA)/SQRT(VA)
      RETURN
      END

C                   ******************
C                   *                *
C                   *   8.  RUNTES   *
C                   *                *
C                   ******************
```

```
      REAL X(1000)
      OPEN(1,FILE=' ')
      WRITE(*,103)
103   FORMAT(2X,'NUMBER OF THE DATA '
     * ' N=?')
      READ(*,*)N
      WRITE(*,104)
104   FORMAT(2X,'THRESHOLD  E=?')
      READ(*,*)E
      READ(1,*)(X(I),I=1,N)
      CALL RUNTES(N,N1,N2,X,IR,E,SIGM,Z)
      CLOSE(1)
      STOP
      END
C*************************************************
      SUBROUTINE RUNTES(N,N1,N2,X,IR,E
     * ,SIGM,Z)
      REAL X(N)
      N1=0
      IR=0
      II=1
      IF(X(1).GT.E)II=-1
      DO 10 K-1,N
      XK=X(K)
      II1=1
      IF(XK.LE.E)II1=-1
      IF(II.EQ.II1)GOTO 11
      IR=IR+1
      II=II1
11    IF(XK-E)12,12,10
12    N1=N1+1
10    CONTINUE
      N2=N-N1
      TN=FLOAT(N)
      T=FLOAT(2*N1*N2)/TN
      E=T+1.
      SIGM=SQRT(T*(T-1.)/(TN-1.))
      Z=(FLOAT(IR)-E)/SIGM
      WRITE(*,101)N1,N2,IR
      WRITE(*,102)E,SIGM,Z
101   FORMAT(2X,3HN1=,I3,3X,3HN2=,I3,3X,3HIR=,I3)
102   FORMAT(2X,'AVERAGE=' G10.4,3X,
     * 'VARIANCE=',G10.4,/,3X,'Z=',G10.4)
```

```
      RETURN
      END

C                  *******************
C                  *                 *
C                  *    9.  GREENF   *
C                  *                 *
C                  *******************

      REAL G(500),F(50),C(50)
      WRITE(*,101)
101   FORMAT(1X,'AR- AND MA- ORDER OF '
     * ,'MODEL IP=?,IQ=?')
      READ(*,*)IP,IQ
      IF(IP.EQ.0)GOTO 10
      WRITE(*,102)IP
102   FORMAT(1X,'AR-PARAMETERS
     * ,'F(I),I=1,....,',I1)
      READ(*,*) (F(I),I=1,IP)
10    IF(IQ.EQ.0)GOTO 11
      WRITE(*,103)IQ
103   FORMAT(1X,'MA-PARAMETERS   ',
     * 'C(I),I=1,....,',I1)
      READ(*,*) (C(I),I=1,IQ)
11    WRITE(*,104)
104   FORMAT(1X,'STEPS OF GREEN FUNCTION',
     * '   L=?')
      READ(*,*)L
      MPQL=MAX0(IP,IQ,L)
      CALL GREENF(IP,IQ,MPQL,L,F,C,G)
      STOP
      END
C***********************************************
      SUBROUTINE GREENF(IP,IQ,MPQL,L,F,C,G)
      REAL F(MPQL),C(MPQL),G(MPQL)
      MP=MAX0(IP,L)
      WRITE(*,102)IP,IQ
      IF(IP.NE.0)WRITE(*,103) (F(I),I=1,IP)
      IF(IQ.NE.0)WRITE(*,104) (C(I),I=1,IQ)
      DO 10 K=1,MPQL
10    G(K)=0.
```

```
      G(1)=1.
      DO 11 J=1,L
      T=0.
      DO 12 K=1,MP
      K1=MP-K+1
      IF(K1.GT.IP)GOTO 13
      T=T+G(K1)*F(K1)
13    IF(K1.EQ.1)GOTO 12
      G(K1)=G(K1-1)
12    CONTINUE
      T1=0.
      IF(J.LE.IQ)T1=C(J)
11    G(1)=T-T1
      T=1.
      WRITE(*,101)T
      DO 14 K=1,MP
      K1=MP-K+1
      IF(K.GT.L)GOTO 14
      WRITE(*,100)K,G(K1)
14    CONTINUE
100   FORMAT(3X,2HG(,I1,2H)=,E14.4)
101   FORMAT(/,8X,'THE GREEN FUNCTION IS:'
     * ,//,3X,5HG(0)=,F14.4)
102   FORMAT(/,5X,5HARMA(,I1,1H,I1,1H),/)
103   FORMAT(16X,2HF=,10F8.3)
104   FORMAT(16X,2HC=,10F8.3)
      RETURN
      END


C                   ******************
C                   *                *
C                   *  10.   INF     *
C                   *                *
C                   ******************


      REAL I(500),F(50),C(50)
      WRITE(*,101)
101   FORMAT(1X,'AR- AND MA- ORDER OF'
     * ,' MODEL  IP=?,  IQ=?')
      READ(*,*)IP,IQ
      IF(IP.EQ.0)GOTO 10
```

```
          WRITE(*,102)IP
102       FORMAT(1X,'AR-PARAMETERS '
     *    ,'F(I),I=1,...,',I1)
          READ(*,*)(F(J),J=1,IP)
10        IF(IQ.EQ.0)GOTO 11
          WRITE(*,103)IQ
103       FORMAT(1X,'MA-PARAMETERS ',
     *    'C(I),I=1,...,',I1)
          READ(*,*)(C(J),J=1,IQ)
11        WRITE(*,104)
104       FORMAT(1X,'STEPS OF INVERSE FUNCTION',
     *    '  L=?')
          READ(*,*)L
          MPQL=MAX0(IP,IQ,L)
          CALL INF(IP,IQ,MPQL,L,F,C,I)
          STOP
          END
C*********************************************
          SUBROUTINE INF(IP,IQ,MPQL,L,F,C,I)
          REAL F(MPQL),C(MPQL),I(MPQL)
          WRITE(*,102)IP,IQ
          IF(IP.NE.0)WRITE(*,103)(F(J),J=1,IP)
          IF(IQ.NE.0)WRITE(*,104)(C(J),J=1,IQ)
          DO 10 K=1,MPQL
10        I(K)=0.
          I(1)=-1.
          DO 11 J=1,L
          T=0.
          DO 12 K=1,MPQL
          K1=MPQL-K+1
          IF(K1.GT.IQ)GOTO 13
          T=T+I(K1)*C(K1)
13        IF(K1.EQ.1)GOTO 12
          I(K1)=I(K1-1)
12        CONTINUE
          T1=0.
          IF(J.LE.IP)T1=F(J)
11        I(1)=T+T1
          T=-1.
          WRITE(*,101)T
          DO 14 K=1,MPQL
          K1=MPQL-K+1
          IF(K.GT.L)GOTO 14
```

```
      WRITE(*,100)K,I(K1)
14    CONTINUE
100   FORMAT(3X,2HI(,I1,2H)=,E14.4)
101   FORMAT(/,8X,'THE INVERSE FUNCTION IS:
     *  ,//,3X,5HI(0)=,E14.4)
102   FORMAT(/,5X,5HARMA(,I1,1H,,I1,1H),/)
103   FORMAT(16X,2HF=,10F8.3)
104   FORMAT(16X,2HC=,10F8.3)
      RETURN
      END

C                 *****************
C                 *               *
C                 *   11. JURYC   *
C                 *               *
C                 *****************

      REAL E(50),F(50)
      WRITE(*,101)
101   FORMAT(2X,'AR/MA - ORDER OF MODEL  N=?')
      READ(*,*)N
      WRITE(*,102)N
102   FORMAT(2X,'AR/MA - PARAMETERS',1X,
     * 'OF MODEL F(I),I=1,...,',I2,' :')
      READ(*,*)(F(I),I=1,N)
      I=JURYC(N,F,E)
      STOP
      END
C**********************************************
      FUNCTION JURYC(N,F,E)
      REAL E(N),F(N)
      E0=-F(N)
      FA=E0
      FB=E0
      TJ=-1.
      DO 11 I=1,N
      I1=N-I
      EI=1.
      IF(I1.GT.0)EI=-F(I1
      E(I)=EI
```

```
        FA=FA+EI
        FB=FB+TJ*EI
11      TJ=-TJ
        IF(FA.LE.0.)GOTO 1010
        Y=FLOAT(N)/2.-FLOAT(N/2)
        IF(Y.EQ.0.)GOTO 14
        IF(FB.GE.0.)GOTO 1010
        GOTO 141
14      IF(FB.LE.0.)GOTO 1010
141     IF(ABS(E0).GE.ABS(E(N)))GOTO 1010
        IF(N.LT.3)GOTO 1012
        J=N-1
1011    EJ=E(J+1)
        JJ=J/2
        TJ=FLOAT(J)/2.-FLOAT(JJ)
        DO 16 I=1,JJ
        I1=J-I+1
        T=E(I)
        T1=E(I1)
        E(I)=E0*T-T1*EJ
16      E(I1)=E0*T1-T*EJ
        IF(TJ)20,21,20
20      JJ=JJ+1
        T=E(JJ)
        E(JJ)=E0*T-T*EJ
21      E0=E0*E0-EJ*EJ
        IF(ABS(E0).LE.ABS(E(J)))GOTO 1010
        IF(J.LE.2)GOTO 1012
        J=J-1
        GOTO 1011
1012    JURYC=0
        WRITE(*,100)JURYC
        WRITE(*,101)
        RETURN
1010    JURYC=1
        WRITE(*,100)JURYC
        WRITE(*,102)
        RETURN
100     FORMAT(/,5X,'JURYC=',I1,/)
101     FORMAT(2X,'ALL ROOTS ARE OUTSIDE'
     *  ,' THE UNIT CIRCLE(STABLE)')
102     FORMAT(2X,'AT LEAST A ROOT IS INSIDE
```

```
     *   ,' THE UNIT CIRCLE(UNSTABLE)')
104      FORMAT(2X,'THE PARAMETERS OF '
     *   ,'ARMA MODEL F=',6F10.2)
         END

C                 ******************
C                 *                *
C                 *  12.  AUTOCR   *
C                 *                *
C                 ******************

         REAL C(50),F(50),RR(200),G(50),B(50,50)
         WRITE(*,101)
101      FORMAT(2X,'AR- AND MA- ORDER  ',
     *   'OF THE MODEL  IP=? , IQ=?')
         READ(*,*)IP,IQ
         IF(IP.EQ.0)GOTO 10
         WRITE(*,103)
103      FORMAT(2X,'AR- PARAMETERS OF MODEL ',
     *   ' F(I),I=1,...,IP  :')
         READ(*,*)(F(I),I=1,IP)
10       IF(IQ.EQ.0)GOTO 11
         WRITE(*,104)
104      FORMAT(2X,'MA- PARAMETERS OF MODEL ',
     *   'C(I),I=1,...,IQ  :')
         READ(*,*)(C(I),I=1,IQ)
11       MPQ1=MAX0(IP,IQ+1)
         WRITE(*,105)
105      FORMAT(2X,'MEAN SQUARE OF THE',
     *   ' SEQUENCE  SIGM=?')
         READ(*,*)SIGM
         WRITE(*,106)
106      FORMAT(2X,'STEPS OF CORRELATION  ',
     *   'FUNCTION  M=?'
         READ(*,*)M
C        EPS  IS COMPUTER ZERO
         EPS=1.E-8
         CALL AUTOCR(M,MPQ1,IP,IQ,F,C,B,G,RR,
     *   R0,SIGM,EPS,ISW)
         STOP
         END
```

```
C***************************************************************
      SUBROUTINE AUTOCR(M,MPQ1,IP,IQ,F,C,B,
     *  G,RR,RO,SIGM,EPS,ISW)
      REAL F(MPQ1),C(MPQ1),RR(M),G(MPQ1)
     *  ,B(MPQ1,MPQ1)
      IF(IP.NE.0)GOTO 1000
      RO=1.
      DO 50 I=1,IQ
50    RR(I)=0.
      DO 54 I=1,IQ
      CI=C(I)
      RO=RO+CI*CI
      IF(I.EQ.IQ)GOTO 54
      I1=IQ-I
      DO 51 J=1,I1
      J1=I+J
      RR(J)=RR(J)+CI*C(J1)
51    CONTINUE
54    CONTINUE
      DO 53 I=1,IQ
53    RR(I)=(-C(I)+RR(I))/RO
      RO=RO/SIGM
      WRITE(*,501)IQ,(C(I),I=1,IQ)
      WRITE(*,505)RO
      WRITE(*,502)
      DO 57 I=1,IQ
57    WRITE(*,503)I,RR(I)
      WRITE(*,504)IQ
501   FORMAT(3X,'MA('I1,') MODEL',1H',
     *  'S PARAMETERS ARE:',/,8X,2HC=,10F8.2)
502   FORMAT(/,4X,'CORRELATION VALUES OF '
     *  'THE MODEL ARE :')
503   FORMAT(5X,3HRR(,I1,2H)=,F12.5)
504   FORMAT(5X,'RR(K)=       0,',6X,
     *  'FOR K>',I1)
505   FORMAT(5X,'VARIANCE OF THE MODEL'
     *  ,'  RO=',F12.5)
      RETURN
1000  MPQ=MPQ1-1
      CALL GREENF(IP,IQ,MPQ1,MPQ1,F,C,G)
      DO 1 I=1,MPQ
      CC=0.
```

```
      IF(I.LE.IQ)CC=-C(I)
1     RR(I+1)=CC
      RR(1)=1.
      DO 12 I=2,MPQ1
      I1=MPQ1-I+2
      GI=G(I1)
      DO 11 J=1,MPQ1
      J1=I+J-2
      J2=J-I
      F2=0.
      IF(J.LE.IQ.AND.J1.LE.IQ)RR(J)=
     *      RR(J)-C(J1)*GI
      IF(J1.LE.IP)F2=F(J1)
      IF(J2)20,21,22
20    F1=0.
      GOTO 11
21    F1=-1.
      GOTO 11
22    F1=F(J2)
11    B(J,I)=-F1-F2
12    B(I,1)=B(1,I)
13    B(1,1)=1.
      DO 14 I=1,MPQ1
14    RR(I)=SIGM*RR(I)
      CALL GAUSS(MPQ1,MPQ1,B,RR,EPS,ISW)
      IF(ISW.EQ.0)GOTO 25
      RO=RR(1)
      WRITE(*,101)IP,IQ,(F(I),I=1,IP)
      IF(IQ.NE.0)WRITE(*,100)(C(I),I=1,IQ)
      WRITE(*,104)RO
      WRITE(*,102)
      DO 2 I=2,MPQ1
      I1=I-1
      RR(I1)=RR(I)/RO
      WRITE(*,105)I1,RR(I1)
2     CONTINUE
      K=MPQ1
18    IF(K.GT.M)GOTO 19
      S=0.
      DO 15 I=1,IP
      FI=F(I)
      Ii=K-I
```

```
15        S=S+FI*RR(I1)
          RR(K)=S
          WRITE(*,105)K,S
          K=K+1
          GOTO 18
19        RETURN
25        WRITE(*,108)ISW
          RETURN
100       FORMAT(8X,2HC=,10F8.2,/)
101       FORMAT(3X,'PARAMETERS OF ARMA('
     *    ,I1,1H,,I1,1H),/8X,2HF=,10F8.2)
102       FORMAT(4X,'CORRELATION VALUES OF'
     *    ' THE MODEL ARE :')
104       FORMAT(6X,'VARIANCE OF THE MODEL '
     *    'R0=',F12.5)
105       FORMAT(5X,'RR(',I1,2H)=,F12.5)
108       FORMAT(5X,4HISW=,I1)
          END
C**************************************************
          SUBROUTINE GAUSS(M,N,A,B,EPS,ISW)
          REAL A(M,M),B(N)
          NM1=N-1
          IF(N.EQ.1)GOTO 30
          DO 10 K=1,NM1
          C=0.
          DO 2 I=K,N
          IF(ABS(A(I,K)).LE.ABS(C))GOTO 2
          C=A(I,K)
          IO=I
2         CONTINUE
          IF(ABS(C).GE.EPS)GOTO 3
          ISW=0
          RETURN
3         IF(IO.EQ.K)GOTO 6
          DO 4 J=K,N
          T=A(K,J)
          A(K,J)=A(IO,J)
4         A(IO,J)=T
          T=B(K)
          B(K)=B(IO)
          B(IO)=T
6         KP1=K+1
```

```
      C=1./C
      B(K)=B(K)*C
      DO 10 J=KP1,N
      A(K,J)=A(K,J)*C
      DO 20 I=KP1,N
20    A(I,J)=A(I,J)-A(I,K)*A(K,J)
10    B(J)=B(J)-A(J,K)*B(K)
30    B(N)=B(N)/A(N,N)
      IF(N.EQ.1)GOTO 31
      DO 40 K=1,NM1
      I=N-K
      C=0.
      IP1=I+1
      DO 50 J=IP1,N
50    C=C+A(I,J)*B(J)
40    B(I)=B(I)-C
31    ISW=1
      RETURN
      END

C                  ******************
C                  *                *
C                  *   13.  PARCR   *
C                  *                *
C                  ******************

      REAL FF(1000),RR(1000),FFK(1000)
      OPEN(1,FILE=' ')
      WRITE(*,106)
106   FORMAT(2X,'NUMBER OF AUTO- AND'/,'
     * ' PARTIAL CORRELATION FUNCTIONS M=?')
      READ(*,*)M
      READ(1,*)(RR(I),I=1,M)
      CALL PARCR(M,FF,RR,FFK)
      CLOSE(1)
      STOP
      END
C*****************************************************
      SUBROUTINE PARCR(M,FF,RR,FFK)
      REAL RR(M),FF(M),FFK(M)
```

```
      FF(1)=RR(1)
      DO 21 K=1,M
      WRITE(*,105)K
      WRITE(*,100)(FF(I),I=1,K)
      FFK(K)=FF(K)
      IF(K.EQ.M)GOTO 21
      T1=0.
      T2=0.
      DO 22 J=1,K
      T3=FF(J)
      J1=K-J+1
      T1=T1+RR(J1)*T3
22    T2=T2+RR(J)*T3
      T3=(RR(K+1)-T1)/(1.-T2)
      FF(K+1)=T3
      I=K/2
      IF(I.EQ.0)GOTO 5
      DO 23 J=1,I
      T1=FF(J)
      J1=K+1-J
      T2=FF(J1)
      FF(J)=T1-T3*T2
23    FF(J1)=T2-T3*T1
5     J=I+1
      IF(I*2-K.NE.0)FF(J)-FF(J)-T3*FF(J)
21    CONTINUE
      WRITE(*,101)M
      DO 10 I=1,M
10    WRITE(*,102)I,I,FFK(I)
105   FORMAT(6X,2HK=,I4)
101   FORMAT(/,3X,I1,'-ORDER PARTIAL'
     *,1X,'CORRELATION IS:')
102   FORMAT(6X,'FF(',I1,1H,I1,2H)=,F12.4)
100   FORMAT(6X,12F10.4)
      RETURN
      END

C                 ******************
C                 *                *
C                 *   14. ARMASP   *
C                 *                *
C                 ******************
```

```
      REAL RUP,IUP,RDWN,IDWN
      DIMENSION PHI(20),THETA(20),X(600)
      DIMENSION YS(600)
      WRITE(*,100)
100   FORMAT('  AR ORDER--- n = ?',/,
     *  '  MA ORDER --- m = ?',/,
     *  '  MIN. FREQ. OF PSD---Fmin=?',/,
     *  '  MAX. FREQ. OF PSD---Fmax=?',/,
     *  '  POINTS OF PSD BETWEEN XMIN AND XMAX'
     *  ' --- N = ?')
      READ(*,*) NAR,NMA,XMIN,XMAX,N
      WRITE(*,200)
200   FORMAT('  VARIANCE OF Xt = ?',/,
     *  '  VARIANCE OF RESIDUAL Et = ?',/,
     *  '  SAMPLING INTERVAL---DEL=?')
      READ(*,*) GAMAZ,SIGZ,DEL
      IF(NAR.NE.0) WRITE(*,300)
300   FORMAT('  AR PARAMETERS---PH1, PH2,....
     +     ',PHn = ?')
      IF(NAR.NE.0) READ(*,*) (PHI(I),I=1,NAR)
      IF(NMA.NE.0) WRITE(*,400)
400   FORMAT('  MA PARAMETERS---THETA1, '
     +     ' THETA2,...,THETAm = ?')
      IF(NMA.NE.0) READ(*,*)(THETA(I),I=1,NMA)
      PI=4.*ATAN(1.)
5     FNYQ=1./(2.*DEL)
      IF(XMAX.GT.FNYQ) XMAX=FNYQ
      XIN=(XMAX-XMIN)/FLOAT(N-1)
      DO 50 M=1,N
      X(M)=XMIN+XIN*FLOAT(M-1)
      RDWN=1.
      IDWN=0.
      IF(NAR.EQ.0) GO TO 15
C   COMPUTE DENOMINATOR OF SPECTRAL DENSITY
      DO 10 I=1,NAR
      CAI=-2.*PI*DEL*I*X(M)
      RDWN=RDWN-PHI(I)*COS(CAI)
10    IDWN=IDWN-PHI(I)*SIN(CAI)
15    RUP=1.
      IUP=0.
C   COMPUTE NUMERATOR OF SPECTRAL DENSITY
      IF(NMA.EQ.0) GO TO 30
      DO 20 J=1,NMA
```

```
        CAI=-2.*PI*DEL*J*X(M)
        RUP=RUP-THETA(J)*COS(CAI)
20      IUP=IUP-THETA(J)*SIN(CAI)
30      CAR=2.*SIGZ*DEL/GAMAZ
        SPNORM=(RUP*RUP+IUP*IUP)*CAR
        SPNORM=SPNORM/(RDWN*RDWN+IDWN*IDWN)
        YS(M)=20.0*ALOG10(AMAX1(SPNORM,1E-20))
50      CONTINUE
        WRITE(*,95)
95      FORMAT(/,1X,'NORMALIZED LOG POWER',
     +  1X,'SPECTRAL DENSITY(db)'/)
        WRITE(*,60)
60      FORMAT(5X,'FREQ',7X,'DB',5X,'FREQ',7X,
     +  'DB',5X,'FREQ',7X,'DB',5X,'FREQ',7X,'DB')
        WRITE(*,70) (X(M),YS(M),M=1,N)
70      FORMAT(4(F9.3,F9.4))
        WRITE(*,80)
80      FORMAT(18H PLOT PSD (1=YES)?)
        READ(*,*) IP
        IF(IP.NE.1) GOTO 90
        CALL GRAPH(YS,N)
90      STOP
        END
C****************************************************
        SUBROUTINE GRAPH(X,N)
        DIMENSION X(N)
        CHARACTER D,BK,DI,B(80)
        DATA D/'*'/,BK/' '/,DI/'.'/
        V=0.0
        DO 1 J=1,N
1       IF(ABS(V).LT.ABS(X(J))) V=X(J)
        WRITE(*,10) D,D,D,V
10      FORMAT(27X,3A1,'MAX=',E14.7)
        WRITE(*,20)
20      FORMAT(28X,80('.'))
        DO 3 I=1,N
           DO 2 II=1,80
2           B(II)=BK
          B(41)=DI
          K=41
          IF(ABS(V).NE.0.) K=X(I)/ABS(V)*39.+41.5
          B(K)=D
          WRITE(*,30) X(I),B
```

```
 30      FORMAT(2X,E14.7,10X,80A1)
 3       CONTINUE
         RETURN
         END
```

```
C                 ******************
C                 *                *
C                 *  15.  CORELA   *
C                 *                *
C                 ******************

        DIMENSION XA(512),XD(1024),YD(1024)
        DIMENSION XC(512),XS(512),ZC(257)
        DIMENSION ZS(257),ILAG(257)
3       OPEN(1,FILE='  ')
        OPEN(2,FILE='  ')
        MAXM=512
        MAXH=MAXM/2+1
C       FILL LAG ARRAY FOR PRINTING
        DO 10 I=1,MAXH
        ILAG(I)=I-1
10      CONTINUE
C       READ IN ANALYSIS PARAMETERS M,N
20      WRITE(*,999)
999     FORMAT(18H SECTION SIZE---M=)
        READ(*,*) M
        IF(M.GT.0.AND.M.LE.MAXM)GOTO 30
        WRITE(*,997)
997     FORMAT(' ILLEGAL INPUT---REENTER VALUE')
        GOTO 20
30      WRITE(*,996)
996     FORMAT(' TOTAL NO. OF DATA---N=')
        READ(*,*) N
        LSHFT=M/2
        MHLF1=LSHFT+1
```

```
      NSECT=(FLOAT(N+LSHFT)-1.)/FLOAT(LSHFT)
C     READ IN MODE DATA TYPE FORMAT
      WRITE(*,994)
994   FORMAT(6H MODE=/
     1 ,10X, 'MODE = 0 AUTO CORRELATION WITHOUT'
     + ,' MEANS SUBTRACTED'/
     2 ,10X,'MODE = 1 CROSS CORRELATION WITHOUT'
     + ,' MEANS SUBTRACTED'/
     3 ,10X,'MODE = 2 AUTO CORRELATION WITH ',
     + 'MEANS SUBTRACTED'/
     4 ,10X,'MODE = 3 CROSS CORRELATION WITH ',
     + 'MEANS SUBTRACTED')
      READ(*,*) MODE
      WRITE(*,990)M,N,MODE
990   FORMAT(3H M=,I4,4H  N=,I5,6H MODE=,I1)
      IS=1
      READ(1,*) (XD(I),I=1,N)
      READ(2,*) (YD(I),I=1,N)
1001  IF(MODE.LT.2)GOTO 80
      NRD=LSHFT
      XSUM=0.
      YSUM=0.
C     LOOP TO CALCULATE MEANS OF X AND Y DATA
      DO 70 K=1,NSECT
      IF(K.EQ.NSECT)NRD=N-(K-1)*NRD
      CALL GETD(XA,XD,NRD,IS)
      IS=IS-NRD
      DO 40 I=1,NRD
      XSUM=XSUM+XA(I)
40    CONTINUE
      IF(MODE.EQ.2)GOTO 60
      CALL GETD(XA,YD,NRD,IS)
      IS=IS-NRD
      DO 50 I=1,NRD
      YSUM=YSUM+XA(I)
50    CONTINUE
60    IS=IS+NRD
70    CONTINUE
      XMEAN=XSUM/FLOAT(N)
      YMEAN=YSUM/FLOAT(N)
      IF(MODE.EQ.2)YMEAN=XMEAN
      XMNC=XMEAN
```

```
      XMNS=YMEAN
      WRITE(*,988)XMEAN,YMEAN
988   FORMAT(7H XMEAN=,E14.5,8H   YMEAN=,E14.5)
C     LOOP TO ACCUMULATE CORRELATIONS
80    IS=1
      NRDY=M
      NRDX=LSHFT
      DO 90 I=1,MHLF1
      ZC(I)=0.
      ZS(I)=0.
90    CONTINUE
      DO 190 K=1,NSECT
      NSECT1=NSECT-1
      IF(K.LT.NSECT1)GOTO 110
      NRDY=N-(K-1)*LSHFT
      IF(K.EQ.NSECT)NRDX=NRDY
      IF(NRDY.EQ.M)GOTO 110
      NRDY1=NRDY+1
      DO 100 I=NRDY1,M
      XC(I)=0.
      XS(I)=0.
100   CONTINUE
110   CALL GETD(XA,XD,NRDY,IS)
      IS=IS-NRDY
      DO 120 I=1,NRDY
      XC(I)=XA(I)
      XS(I)=XA(I)
120   CONTINUE
      IF(MODE.EQ.0.OR.MODE.EQ.2)GOTO 140
      CALL GETD(XA,YD,NRDY,IS)
      IS=IS-NRDY
      DO 130 I=1,NRDY
      XS(I)=XA(I)
130   CONTINUE
140   IF(MODE.LT.2)GOTO 160
      DO 150 I=1,NRDY
      XC(I)=XC(I)-XMNC
      XS(I)=XS(I)-XMNS
150   CONTINUE
160   NRDX1=NRDX+1
      DO 170 I=NRDX1,M
      XC(I)=0.
```

```
 170     CONTINUE
C        CORRELATE X AND Y SECTIONS
         CALL FFT(XC,XS,M,0)
         DO 180 I=2,LSHFT
         J=M+2-I
         XIC=(XC(I)+XC(J))*0.5
         XIS=(XS(I)-XS(J))*0.5
         YIC=(XS(I)+XS(J))*0.5
         YIS=(XC(J)-XC(I))*0.5
         ZC(I)=ZC(I)+XIC*YIC+XIS*YIS
         ZS(I)=ZS(I)+XIC*YIS-XIS*YIC
 180     CONTINUE
         ZC(1)=ZC(1)+XC(1)*XS(1)
         ZC(MHLF1)=ZC(MHLF1)+XC(MHLF1)*XS(MHLF1)
         IS=IS+LSHFT
 190     CONTINUE
C        INVERSE DFT TO GIVE CORRELATION
         DO 200 I=2,LSHFT
         J=M+2-I
         XC(I)=ZC(I)
         XS(I)=ZS(I)
         XC(J)=ZC(I)
         XS(J)=-ZS(I)
 200     CONTINUE
         XC(1)=ZC(1)
         XS(1)=ZS(1)
         XC(MHLF1)=ZC(MHLF1)
         XS(MHLF1)=ZS(MHLF1)
         CALL FFT(XC,XS,M,1)
         FN=FLOAT(N)
         DO 210 I=1,MHLF1
         XA(I)=XC(I)/FN
 210     CONTINUE
C        PLOT AND PRINT THE CORRELATION FUNCTION
C        WHICH IS IN THE ARRAY XA
         WRITE(*,205)
 205     FORMAT(' PLOT THE CORRECATION FUNCTION
     +    ,'(1=YES,0=NO)=')
         READ(*,*) IE
         IF(IE.EQ.0) GOTO 9987
         CALL GRAPH(XA,MHLF1)
 9987    WRITE(*,987)
```

```
987       FORMAT(21H CORRELATION FUNCTION)
          WRITE(*,986)
986       FORMAT(1X,3HLAG,2X,4HCORR,5X,3HLAG,2X,
     *    4HCORR,5X,3HLAG,2X,4HCORR,5X,3HLAG,2X,
     *    4HCORR,5X,3HLAG,2X,4HCORR)
          WRITE(*,985)(ILAG(I),XA(I),I=1,MHLF1)
985       FORMAT(5(I4,E10.3))
          CLOSE(1)
          CLOSE(2)
          STOP
          END
C*************************************************************
C         READ REAL DATA ROUTINE
          SUBROUTINE GETD(X,XD,NRD,IS)
          DIMENSION X(1),XD(1)
          DO 10 I=1,NRD
          X(I)=XD(IS)
          IS=IS+1
10        CONTINUE
          RETURN
          END
C*************************************************************
          SUBROUTINE FFT(XC,XS,N,INV)
          DIMENSION XC(1024),XS(1024)
          REAL UC,US,WC,WS,TC,TS
          M=ALOG(FLOAT(N))/ALOG(2.)+0.1
          NV2=N/2
          NM1=N-1
          J=1
          DO 40 I=1,NM1
          IF(I.GE.J)GOTO 10
          TC=XC(J)
          TS=XS(J)
          XC(J)=XC(I)
          XS(J)=XS(I)
          XC(I)=TC
          XS(I)=TS
10        K=NV2
20        IF(K.GE.J)GOTO 30
          J=J-K
          K=K/2
          GOTO 20
30        J=J+K
```

```
40        CONTINUE
          PI=4.*ATAN(1.0)
          DO 70 L=1,M
          LE=2**L
          LE1=LE/2
          UC=1.
          US=0.
          WC=COS(PI/FLOAT(LE1))
          WS=-SIN(PI/FLOAT(LE1))
          IF(INV.NE.0)WS=-WS
          DO 60 J=1,LE1
          DO 50 I=J,N,LE
          IP=I+LE1
          TC=XC(IP)*UC-XS(IP)*US
          TS=XS(IP)*UC+XC(IP)*US
          XC(IP)=XC(I)-TC
          XS(IP)=XS(I)-TS
          XC(I)=XC(I)+TC
          XS(I)=XS(I)+TS
50        CONTINUE
          UC1=UC*WC-US*WS
          US=US*WC+UC*WS
          UC=UC1
60        CONTINUE
70        CONTINUE
          IF(INV.EQ.0)RETURN
          DO 80 I=1,N
          XC(I)=XC(I)/FLOAT(N)
          XS(I)=XS(I)/FLOAT(N)
80        CONTINUE
          RETURN
          END

C                    ******************
C                    *                *
C                    *  16.   INGUS   *
C                    *                *
C                    ******************

          REAL B(50),D(50),E(50),ZR(50),ZI(50),
```

```
     *   R(50),RB(50),F(50),C(50),EF(50,50),A(50)
         OPEN(1,FILE=' ')
         WRITE(*,101)
101      FORMAT(2X,'AR- AND MA- ORDER OF'
     *   ' THE MODEL  IP=?  IQ=?')
         READ(*,*)IP,IQ
C        EPS  IS COMPUTER ZERO
         EPS=1.E-10
         IPST=MAXO(IP,IQ)+IQ
         READ(1,*)(R(I),I=1,IPST)
C        KK IS MAXIMUM NUMBER OF ITERATION FOR
C        SOLVING EQUATION
         KK=8
         IQ1=IQ+1
         CALL ARMAI(IP,IQ,IPST,RB,EE,F,C,
     *   R,EPS,ISW)
         IF(IQ.EQ.0)GOTO 999
         CQ=C(IQ)
         DO 10 I=1,IQ
         I1=IQ-1+1
10       A(I)=C(I1)/CQ
         A(IQ1)=-1./CQ
         CALL ROOTI(IQ1,A,ZR,ZI,B,D,E,KK)
         AQ1=A(IQ+1)
         DO 11 I=1,IQ
         I1=IQ-I+1
11       C(I1)=-A(I)/AQ1
         WRITE(*,103)
         WRITE(*,104)(C(I),I=1,IQ)
103      FORMAT(2X,'*** ESTIMATION OF ',
     *   'MA- PARAMETERS  ARE:')
104      FORMAT(10F8.3)
999      STOP
         END
C***********************************************************
         SUBROUTINE ARMAI(IP,IQ,IPST,FF,EE,
     *   F,C,RR,EPS,ISW)
         REAL F(IPST),C(IPST),FF(IPST),
     *   EE(IPST,IPST),RR(IPST)
         CALL PARCOR(IPST,IPST,FF,RR,F)
         MPQ=MAXO(IP,IQ)
         FMPQ=FF(MPQ)
         IF(IQ-1)2,3,4
```

```
4       DO 1 I=2,IQ
        IO=IQ-I+1
        I1=MPQ-I+1
        I2=MPQ+I-1
        FF1=FF(I1)
        FF2=FF(I2)
        C(I)=FF(I2+1)
        EE(I,I)=FMPQ
        DO 1 J=1,IO
        J1=I+J-1
        EE(J1,J)=FF2
1       EE(J,J1)=FF1
3       EE(1,1)=FMPQ
        C(1)=FF(1PST)
        CALL GAUSS(IPST,IQ,EE,C,EPS,ISW)
        IF(ISW.EQ.1)GOTO 40
        WRITE(*,101)
101     FORMAT(2X,'MA- PARAMETERS :')
        WRITE(*,100) (C(I),I=1,IQ)
100     FORMAT(10F12.4)
        IF(IP.EQ.0)GOTO 40
2       DO 30 J=1,IP
        CO=0.
        IF(J.LE.IQ)CO=C(J)
        T=CO+FF(J)
        JO=J-1
        IF(JO.EQ.0)GOTO 30
        DO 31 I=1,JO
        J1=J-I
        C1=0.
        IF(I.LE.IQ)C1=C(I)
31      T=T-C1*FF(J1)
30      F(J)=T
        WRITE(*,102)
102     FORMAT(2X,'*** ESTIMATION OF '
     *  'AR- PARAMETERS  :')
        WRITE(*,100) (F(I),I=1,IP)
40      RETURN
        END
C*****************************************************************
        SUBROUTINE ROOTI(MM,A,ZR,ZI,B,D,E,KK)
        REAL A(MM),ZR(MM),ZI(MM),B(MM),
     *  D(MM),E(MM)
```

```
      M=MM-1
      DO 7 I=1,MM
7     B(I)=A(I)
      CALL RTPOL(MM,B,ZR,ZI,D,E,100,KK)
      I=1
      B(1)=1.
10    IF(I.GT.M)GOTO 11
      T1=ZR(I)
      T2=ZI(I)
      G2=T1*T1+T2*T2
      IF(G2.LE.1.)GOTO 13
      IF(ABS(T2)-1E-12)14,14,17
13    IF(ABS(T2).LT.1E-12)GOTO 18
      G1=-2.*ZR(I)
      BO=1.
      B1=O.
      M2=M-2
      IF(M2.EQ.O)GOTO 31
      DO 15 L1=1,M2
      L=L1+1
      BB=A(L)-G1*BO-G2*B1
      B(L)=BB
      B1=BO
      BO=BB
15    CONTINUE
      M21=M2+1
      G2=1./G2
      G1=G1*G2
      BO=1.
      B1=O.
      DO 16 L1=1,M2
      L=L1+1
      A(L)=B(L)+G1*BO+G2*B1
      B1=BO
      BO=B(L)
16    CONTINUE
      A(M)=G1*B(MM-2)+G2*B(MM-3)
      A(MM)=G2*B(MM-2)
      GOTO 17
31    A(M)=G1/G2
      A(MM)=1./G2
17    I=I+2
```

```
      GOTO 10
18    B0=1.
      M1=M-1
      IF(M1.LE.0)GOTO 32
      DO 19 L1=1,M1
      L=L1+1
      B(L)=A(L)+B0*T1
      B0=B(L)
19    CONTINUE
      T1=1./T1
      B0=1.
      DO 21 L1=1,M1
      L=L1+1
      A(L)=-B0*T1+B(L)
      B0=B(L)
21    CONTINUE
      A(MM)=-B(M)*T1
      GOTO 14
32    A(MM)=1./T1
14    I=I+1
      GOTO 10
11    CONTINUE
      DO 9 I=1,MM
9     B(I)=A(I)
      WRITE(*,103)
103   FORMAT(2X,'*****************************'
     * ,'***********')
      CALL RTPOL(MM,B,ZR,ZI,D,E,100,KK)
      RETURN
      END
      SUBROUTINE RTPOL(N,A,B,C,D,E,L,K)
      REAL A(N),B(N),C(N),D(N),E(N)
      T=1.
      G=10.**K
      M=N
2     IF(A(M).NE.0.)GOTO 6
      A(M)=G
      B(M)=0.
      C(M)=0.
      M=M-1
      GOTO 2
6     IF(M.LE.3)GOTO (300,90,190),M
      PS=0.
```

```
          QS=0.
          PT=Q.
          QT=0.
          REV=1.
          G=10.**K
          S=0.
          DO 16 J=1,M
          IF(A(J).NE.O.)S=ALOG(ABS(A(J)))+S
16        CONTINUE
          S=EXP(S/FLOAT(M))
          DO 18 J=1,M
18        A(J)=A(J)/S
          IF(ABS(A(2)*A(M))-ABS(A(M-1)*A(1))
     *       )20,22,22
20        T=-1
          MM=M/2
          DO 24 J=1,MM
          S=A(J)
          I=M-J+1
          A(J)=A(I)
24        A(I)=S
22        IF(QS)26,28,26
26        P=PS
          Q=QS
          GOTO 50
28        IF(A(M-2).NE.O.)GOTO 32
          Q=1.
          P=-2.
          GOTO 38
32        Q=A(M)/A(M-2)
          P=(A(M-1)-Q*A(M-3))/A(M-2)
38        R=0.
50        DO 74 I=1,L
          B(1)=A(1)
          C(1)=B(1)
          B(2)=A(2)-P*B(1)
          C(2)=B(2)-P*C(1)
          DO 54 J=3,M
          B(J)=A(J)-P*B(J-1)-Q*B(J-2)
54        C(J)=B(J)-P*C(J-1)-Q*C(J-2)
          IF(B(M-1)*A(M-1))56,58,56
56        IF(ABS(A(M-1)/B(M-1)).LT.G)GOTO 60
          B(M)=A(M)-Q*B(M-2)
```

```
 58       IF(ABS(B(M))*G.LT.ABS(A(M)))GOTO 200
 60       D(1)=A(1)
          E(1)=D(1)
          DO 62 J=2,M
          D(J)=A(J)+R*D(J-1)
 62       E(J)=D(J)+R*E(J-1)
          IF(ABS(D(M))*G.LT.ABS(A(M)))GOTO 100
          C(M-1)=C(M-1)-B(M-1)
          S=C(M-2)**2-C(M-1)*C(M-3)
          IF(S)64,66,64
 66       P=P-2
          Q=Q*(Q+1.)
          GOTO 68
 64       P=(B(M-1)*C(M-2)-B(M)*C(M-3))/S+P
          Q=(B(M)*C(M-2)-B(M-1)*C(M-1))/S+Q
 68       IF(E(M-1))70,72,70
 72       R=R-1.
          GOTO 74
 70       R=R-D(M)/E(M-1)
 74       CONTINUE
          PS=PT
          QS=QT
          PT=P
          QT=Q
          IF(REV.LT.O.)G=G/10.
          REV=-REV
          GOTO 20
 90       R=-A(2)/A(1)
 100      IF(T.LT.O.)R=1./R
          A(M)=S
          B(M)=R
          C(M)=O.
          IF(M.EQ.2)GOTO 300
          M=M-1
          DO 110 J=1,M
          IF(ABS(A(J))-G*ABS(D(J)))120,130,130
 120      A(J)=D(J)
          GOTO 110
 130      A(J)=O.
 110      CONTINUE
          IF(M.LE.3)GOTO 6
          GOTO 38
```

```
190       P=A(2)/A(1)
          Q=A(3)/A(1)
200       IF(T.GE.0.)GOTO 210
          P=P/Q
          Q=1./Q
210       P=-.5*P
          IF(P*P-Q)220,230,230
220       B(M)=P
          B(M-1)=P
          S=SQRT(Q-P*P)
          C(M)=S
          C(M-1)=-S
          GOTO 270
230       S=SQRT(P*P-Q)
          IF(P)240,240,250
240       B(M)=P-S
          GOTO 260
250       B(M)-P+S
260       B(M-1)=Q/B(M)
          C(M)=0.
          C(M 1)=0.
270       A(M)=G
          IF(M.EQ.3)GOTO 300
          M=M-2
          DO 295 J=1,M
          IF(ABS(A(J))-G*ABS(B(J)))280,290,290
280       A(J)=B(J)
          GO(O 295
290       A(J)=0.
295       CONTINUE
          GOTO 6
300       DO 310 J=2,N
          A(J-1)=A(J)
          B(J-1)=B(J)
310       C(J-1)=C(J)
          N1=N-1
          WRITE(*,500)
500       FORMAT(5X,'REAL-PART',5X,'IMAGINARY'
     *    '-PART',5X,'MAGNITUDE',/3X,'OF THE   ',
     *    'ROOT',7X,' THE ROOT',2X,'OF THE ROOT')
          DO 5 I=1,N1
          BT=B(I)
          CT=C(I
```

```
      TT=SQRT(BT*BT+CT*CT)
5     WRITE(*,502)BT,CT,TT
502   FORMAT(3F14.5)
      RETURN
      END

C               ******************
C               *                *
C               *  17.  ARMAME   *
C               *                *
C               ******************

      REAL RR(50),RRB(50),F(50),C(50),CC(50),
     * EE(50,50)
      OPEN(1,FILE=' ')
      WRITE(*,101)
101   FORMAT(2X,'AR- AND MA- ORDER OF THE '
     * 'MODEL  IP=?  IQ=?')
      READ(*,*)IP,IQ
      IPQ=IP+IQ
      READ(1,*)(RR(I),I=1,IPQ)
      WRITE(*,102)
102   FORMAT(2X,'MAXIMUM NUMBER OF ITERATION'
     * ,'  ML=?')
      READ(*,*)ML
      WRITE(*,103)
103   FORMAT(2X,'TULERANCE VALUE  EPSi=? )
      READ(*,*)EPS1
C     EPS2  IS COMPUTER ZERO
      EPS2=1.E-10
      CALL ARMAME(IP,IQ,IPQ,RR,F,C,CC,EE,
     * RRB,EPS1,EPS2,ISW,ML)
      STOP
      END
C*************************************************
      SUBROUTINE ARMAME(IP,IQ,IPQ,RR,F.C,
     * CC,EE,RRB,EPS1,EPS2,ISW,ML)
      REAL RR(IPQ),EE(IPQ,IPQ),RRB(IPQ),
     * F(IPQ),C(IPQ),CC(IPQ)
      IF(IP.EQ 0)GOTO 19
      RRQ=1.
```

```
          IF(IQ.NE.O)RRQ=RR(IQ)
          IF(IP.EQ.1)GOTO 2
          DO 1 I=2,IP
          IO=IP-I+1
          I1=IQ-I+1
          I2=IQ+I-1
          IF(I1)11,12,13
11        I1=-I1
          GOTO 13
12        RR1=1.
          GOTO 14
13        RR1=RR(I1)
14        RR2=RR(I2)
          F(I)=RR(I2+1)
          EE(I,I)=RRQ
          DO 1 J=1,IO
          J1=I+J-1
          EE(J,J1)=RR1
1         EE(J1,J)=RR2
2         EE(1,1)=RRQ
          F(1)=RR(IQ+1)
          CALL GAUSS(IPQ,IP,EE,F,EPS2,ISW)
          IF(ISW.EQ.1)GOTO 40
19        WRITE(*,107)IP,IQ
          IF(IP.NE.O)WRITE(*,106)(F(I),I=1,IP)
106       FORMAT(10F10.3)
107       FORMAT(1X,'ARMA(',I1,1H,,I1,')',
     *    ' PARAMETERS',1H',' ESTIMATION :')
          IF(IQ.EQ.O)GOTO 40
          IQ1=IQ+1
          DO 3 KK=1,IQ1
          K=KK-1
          RK=1.
          IF(K.NE.O)RK=RR(K)
          RRK=RK
          DO 30 I=1,IP
          I1=IP-I
          K1=K+I
          K2=K-I
          IF(K2)32,33,34
32        K2=-K2
          GOTO 34
33        R2=1.
```

```
      GOTO 35
 34   R2=RR(K2)
 35   R1=RR(K1)
      R12=R1+R2
      FI=F(I)
      RRK=RRK-FI*R12+FI*FI*RK
      IF(I1.EQ.0)GOTO 30
      DO 31 J=1,I1
      J1=I+J
      RRK=RRK+F(J)*F(J1)*R12
 31   CONTINUE
 30   CONTINUE
      IF(K)37,37,36
 37   RRB0=RRK
      GOTO 3
 36   RRB(K)=RRK/RRB0
 3    CONTINUE
      CALL MAME(IQ,C,RRB,CC,SIGM,RRB0,
     * EPS1,ML)
 40   RETURN
      END
C***********************************************************
      SUBROUTINE MAME(IQ,C,RR,CC,SIGM,
     * RO,EPS,ML)
      REAL C(IQ),RR(IQ),CC(IQ)
      L=1
      DO 10 I=1,IQ
 10   C(I)=0.
 18   SIGM=1.
      KK=0
      DO 13 I=1,IQ
 13   CC(I)=0.
      DO 11 I=1,IQ
      CI=C(I)
      SIGM=SIGM+CI*CI
      I1=IQ-I
      IF(I.EQ.IQ)GOTO 17
      DO 11 J=1,I1
      J1=J+I
      CC(J)=CC(J)+C(J1)*CI
 11   CONTINUE
 17   SIGM=RO/SIGM
```

```
      DO 12 I=1,IQ
      CC(I)=CC(I)-RR(I)/SIGM*RO
      K=1
      IF(ABS(CC(I)-C(I)).GT.EPS)K=0
      C(I)=CC(I)
      KK=KK*K
12    CONTINUE
      IF(KK.EQ.1.OR.L.GT.ML)GOTO 15
      L=L+1
      GOTO 18
15    WRITE(*,103)(C(I),I=1,IQ)
      IF(KK.GT.1)GOTO 31
      RETURN
31    WRITE(*,107)
      RETURN
103   FORMAT(8E13.4)
107   FORMAT(/5X,'ESTIMATES DO NOT '
     *  ,'MEEDS OF PRECISION')
      END

C                *******************
C                *                 *
C                *   18.   LSME    *
C                *                 *
C                *******************

      SUBROUTINE LSME(N1,N2,MRE,S,
     *  SZ,E,SS,F,FO,A,SO,AIC,ISW)
      REAL S(MRE,MRE),SZ(MRE),E(MRE),F(MRE)
     *  ,A(MRE),SS(MRE)
      NO=0
      EZ=0.
      ZZ=0.
      DO 10 I=1,MRE
      SZ(I)=0.
      SS(I)=0.
      E(I)=0.
      DO 10 J=I,MRE
10    S(I,J)=0.
      DO 11 K=N1,N2
      CALL SUBLSM(NO,MRE,K,A,ZK)
```

```
      EZ=EZ+ZK
      ZZ=ZZ+ZK*ZK
      DO 11 I=1,MRE
      AKI=A(I)
      SZ(I)=SZ(I)+AKI*ZK
      E(I)=E(I)+AKI
      SS(I)=SS(I)+AKI*AKI
      DO 11 J=I,MRE
      AKJ=A(J)
11    S(I,J)=S(I,J)+AKI*AKJ
      TN=FLOAT(N2-N1+NO+1)
      EZ=EZ/TN
      ZZ=SQRT(ZZ-TN*EZ*EZ)
      DO 14 I=1,MRE
      T=E(I)/TN
      SS(I)=SQRT(SS(I)-TN*T*T)
14    E(I)=T
      DO 15 I=1,MRE
      T1=E(I)
      T2=SS(I)
      SZ(I)=(SZ(I)-TN*EZ*T1)/(ZZ*T2)
      DO 15 J=I,MRE
      T=(S(I,J)-TN*T1*E(J))/(SS(J)*T2)
      IF(J.NE.I)S(J,I)=T
15    S(I,J)=T
      CALL LDLT(MRE,S,A,SZ,F,ISW)
      IF(ISW.EQ.1)GOTO 99
      T1=1.0
      FO=EZ
      DO 16 I=1,MRE
      T2=F(I)
      T=T2*ZZ/SS(I)
      T1=T1-SZ(I)*T2
      FO=FO-E(I)*T
16    F(I)=T
      SO=ZZ*ZZ*T1/TN
      AIC=ALOG(SO)+2.*FLOAT(MRE)/TN
      SO=SQRT(SO)
99    RETURN
      END
C**********************************************************
      SUBROUTINE LDLT(N,A,D,B,XX,ISW)
```

```
      REAL  A(N,N),D(N),B(N),XX(N)
      DO 5 I=1,N
      DO 5 J=1,I
      W=A(J,I)
      IF(J-I)10,20,10
20    IF(I.EQ.1)GOTO 30
      IM1=I-1
      DO 25 K=1,IM1
      Y=A(I,K)
      T=Y*D(K)
      A(I,K)=T
25    W=W-Y*T
30    IF(W.NE.0.)GOTO 35
      ISW=1
      WRITE(*,201)ISW
      RETURN
35    D(I)=1.0/W
      GOTO 5
10    IF(J.EQ.1)GOTO 45
      JM1=J-1
      DO 40 K=1,JM1
40    W=W-A(I,K)*A(J,K)
45    A(I,J)=W
5     CONTINUE
      DO 60 I=1,N
      Y=B(I)
      IF(I.EQ.1)GOTO 60
      IM1=I-1
      DO 65 K=1,IM1
65    Y=Y-A(I,K)*XX(K)
60    XX(I)=Y
      DO 100 L=1,N
      I=N-L+1
      Y=XX(I)*D(I)
      IF(I.EQ.N)GOTO 100
      IP1=I+1
      DO 70 K=IP1,N
70    Y=Y-A(K,I)*XX(K)
100   XX(I)=Y
      ISW=0
201   FORMAT(5X,4HISW=,I1)
      RETURN
```

```
      END

C                    ******************
C                    *                *
C                    *  19.   MARQT   *
C                    *                *
C                    ******************

      DOUBLE PRECISION GRD,SIGMA
      DIMENSION PAR(80)
      COMMON/BLOK1/N,M,NOB,NPAR
      COMMON/BLOK3/SIGMA,GRD(40),BNDLW(40),
     +  BNDUP(40),ITMAX,TOL(2),VAR
      COMMON/BLOK4/X(1024),OBS(1024),AT(1024)
      OPEN(1,FILE='  ')
      WRITE(*,1030)
1030  FORMAT(/,' NO. OF DATA---NOB=',/,
     * ' OPTION FOR MEAN ESTIMATION---IMEAN='.
     + '(YES=1,NO=0)')
      READ(*,*) NOB,IMEAN
      READ(1,*) (X(I),I=1,NOB)
      CLOSE(1)
C       NORMALIZE TIME SERIES X(I)
190   SUM=0.
      RNOB=FLOAT(NOB)
      DO 200 I=1,NOB
200   SUM=SUM+X(I)
      AVG=SUM/RNOB
      SUM=0.
      DO 210 I=1,NOB
210   SUM=SUM+(X(I)-AVG)*(X(I)-AVG)
      VAR=SUM/RNOB
      STD=SQRT(VAR)
      DO 230 I=1,NOB
      X(I)=(X(I)-AVG)/STD
230   OBS(I)=X(I)
1000  WRITE(*,1010)
1010  FORMAT(' AR ORDER--N=',/' MA ORDER--M=')
      READ(*,*)N,M
      WRITE(*,1050)
1050  FORMAT(/,' MAX. NO. OF ITERATIONS---',
```

```
     *    'ITMAX=',/ ' TOLERANCE FOR THE CHANGE ',
     *     'OF PARAMETERS---TOL(1)=',/
     +    ' CONVERGENT CRITERION---TOL(2)=')
          READ(*,*)ITMAX,TOL(1),TOL(2)
          NPAR=N+M
          IP=N+M
          IF(IMEAN.GT.0) NPAR=NPAR+1
          IF(IMEAN.GT.0) PAR(NPAR)=0.
          WRITE(*,235)
235       FORMAT(' INPUT INITIAL VALUES OF ',
     +     'PHI(S) AND THETA(S)')
          READ(*,*) (PAR(K),K=1,IP)
          WRITE(*,280)
          CALL MAR(PAR)
C           PRINT OUT THE PARAMETERS OF MODEL
          WRITE(*,280)
280       FORMAT(1X,60(1H-)/)
          WRITE(*,290)N,M
290       FORMAT(1H ,10X,'ARMA(',I2,',',I2,')'/)
          IF(IMEAN.EQ.0) GO TO 312
          XMEAN=PAR(NPAR)*STD+AVG
          XMUP=BNDUP(NPAR)*STD+AVG
          XMLW=BNDLW(NPAR)*STD+AVG
          WRITE(*,300)XMUP
          WRITE(*,310)XMEAN, VAR,AVG
          WRITE(*,300)XMLW
300       FORMAT( 9X,G15.6)
310       FORMAT(3X,'MEAN =',G15.6,3X,'VAR =',
     +     G15.6,3X,'AVG =',G15.6)
          GO TO 314
312       WRITE(*,313)VAR,AVG
313       FORMAT(5X,'VAR=',G15.6,5X,'AVG=',G15.6)
314       IF(N.GT.0)WRITE(*,320)
320       FORMAT(20X,'AUTOREGRESSIVE PARAMETERS')
          NS=N
          ISTART=1
          ISTOP=N-((N-1)/10)*10
          ISTOP=6+(ISTOP-1)/2
          INC=ISTOP
330       DO 340 K=1,10
          IF(ISTOP.GT.NS) ISTOP=NS
          IF(ISTOP.LT.ISTART) GO TO 335
```

```
      WRITE(*,380) (BNDUP(I),I=ISTART,ISTOP)
      WRITE(*,380) (PAR(I),I=ISTART,ISTOP)
      WRITE(*,380) (BNDLW(I),I=ISTART,ISTOP)
335   CONTINUE
      ISTART=ISTOP+1
      ISTOP=ISTOP+INC
      IF(ISTART.GT.NS) GO TO 350
340   CONTINUE
350   IF(ISTART.GT.IP) GOTO 390
      NS=IP
      WRITE(*,360)
360   FORMAT(20X,' MOVING AVERAGE PARAMETERS')
      GO TO 330
380   FORMAT(1X,10F12.6)
390   SIGMA=SIGMA*VAR
      WRITE(*,450) SIGMA
450   FORMAT(/,'      SIGMA E-SQUARE =',G12.5)
      WRITE(*,510)  (GRD(I),I=1,NPAR)
510   FORMAT('    MAGNITUDE OF GRADIENT VECTOR'
     + ,10(/,5X,10F12.5))
      SIG=SNGL(SIGMA)
      AIC=RNOB*ALOG(SIG)+2.*FLOAT(N+M+1)
      BIC=AIC+FLOAT(N+M+1)*(ALOG(RNOB)-1.)
      IF(IMEAN.EQ.0) AIC=AIC-2.
      IF(IMEAN.EQ.0) BIC=AIC+FLOAT(N+M)*
     +     (ALOG(RNOB)-1.)
      WRITE(*,100) AIC,BIC
100   FORMAT('   AIC =',F12.5,10X,
     +  'BIC =',F12.5)
      STOP
      END
C-----------------------------------------------------------------------
C
C                  MARQUARDT COMPROMISE ROUTINE
C
C-----------------------------------------------------------------------
        SUBROUTINE MAR(PAR)
        DOUBLE PRECISION GRD,SIGMA
        DIMENSION S(40),DX(40),A(40,40),PAR(1)
     *  ,XTX(40,40),ZA(1024),ZB(1024),ZC(1024)
        DIMENSION D(40)
        COMMON/BLOK1/NAR,NMA,NOB,NPAR
```

```
      COMMON/BLOK3/SIGMA,GRD(40),BNDLW(40),
     4  BNDUP(40),ITMAX,TOL(2),VAR
      COMMON/BLOK4/X(1024),OBS(1024),AT(1024)
      SPEED=1.8
      AMDA=NPAR*1.5
C.       CALCULATE AT
      CALL CAT(PAR)
      JTR=1
      JTRIG=1
      SSOLD=0.
      DO 5 IOB=1,NOB
5     SSOLD=SSOLD+AT(IOB)**2
C.       NIT=NO. OF ITERATIONS
      NIT=1
10    LGA=0
C.       CALCULATE PARTIAL DERIVATIVES OF AT
      CALL JAC(PAR,ZA,1)
      NM=NAR+1
      CALL JAC(PAR,ZB,NM)
      IF(NPAR-NMA-NAR)20,80,20
20    CALL JAC(PAR,ZC,NPAR)
      UU=0.
      S(NPAR)=0.
      DO 30 J=1,NOB
      S(NPAR)=S(NPAR)-ZC(J)*AT(J)
30    UU=UU+ZC(J)**2
      A(NPAR,NPAR)=UU
      DO 50 K=1,NAR
      UU=0.
      KK=K+1
      DO 40 J=KK,NOB
      JJ=J-K+1
40    UU=UU+ZC(J)*ZA(JJ)
      A(NPAR,K)=UU
50    A(K,NPAR)=UU
      DO 70 K=1,NMA
      UU=0.
      KK=K+1
      DO 60 J=KK,NO.
      JJ=J-K+1
60    UU=UU+ZC(J)*ZB(JJ)
      LL=NAR+K
```

```
      A(NPAR,LL)=UU
70    A(LL,NPAR)=UU
80    DO 100 J=1,NAR
      UU=O.
      KK=J+1
      DO 90 L=KK,NOB
      LL=L-J+1
90    UU=UU-AT(L)*ZA(LL)
100   S(J)=UU
      J=1
      DO 120 K=J,NAR
      UU=O.
      KK=K+1
      DO 110 L=KK,NOB
      LL=L-J+1
      JJ=L-K+1
110   UU=UU+ZA(LL)*ZA(JJ)
      A(J,K)=UU
120   A(K,J)=UU
      DO 130 J=2,NAR
      DO 130 K=J,NAR
      JJ=J-1
      KK=K-1
      UU=A(JJ,KK)
      A(J,K)=UU
130   A(K,J)=UU
      DO 140 J=1,NAR
      DO 140 K=1,NMA
      KK=K+1
      IF(K-J) 132,132,133
132   KK=J+1
133   UU=O.
      DO 135 L=KK,NOB
      LL=L-J+1
      JJ=L-K+1
135   UU=UU+ZA(LL)*ZB(JJ)
      JJ=NAR+K
      A(J,JJ)=UU
140   A(JJ,J)=UU
      DO 160 J=1,NMA
      UU=O.
      KK=J+1
```

```
      DO 150 L=KK,NOB
      LL=L-J+1
150   UU=UU-AT(L)*ZB(LL)
      LL=NAR+J
160   S(LL)=UU
      J=1
      DO 180 K=J,NMA
      UU=0.
      KK=K+1
      DO 170 L=KK,NOB
      LL=L-J+1
      JJ=L-K+1
170   UU=UU+ZB(LL)*ZB(JJ)
      LL=NAR+K
      JJ=NAR+J
      A(LL,JJ)=UU
180   A(JJ,LL)=UU
      IF(NMA-1)220,220,190
190   DO 210 J=2,NMA
      DO 200 K=J,NMA
      KK=K-1+NAR
      JJ=J-1+NAR
      UU=A(JJ,KK)
      KK=KK+1
      JJ=JJ+1
      A(KK,JJ)=UU
200   A(JJ,KK)=UU
210   CONTINUE
220   CONTINUE
      DO 230 J=1,NPAR
      DO 225 I=1,NPAR
225   XTX(I,J)=A(I,J)
      GRD(J)=2.*S(J)
230   A(J,J)=A(J,J)+AMDA
C              TO GET INCREMENT OF PARAMETERS
      CALL EQN(A,NPAR,DX,S)
      DO 250 J=1,NPAR
      PAR(J)=PAR(J)+DX(J)
      IF(ABS(DX(J)/PAR(J))-TOL(J))250,250,240
240   LGA=1
250   CONTINUE
      CALL CAT(PAR)
```

```
      SSNEW=0.
      DO 280 I=1,NOB
280   SSNEW=SSNEW+AT(I)**2
      SUMSQ=SSNEW*VAR
      WRITE(*,290)SUMSQ,AMDA
290   FORMAT(' SUM OF SQUARE=      ',E16.8,
     +   3X,'LAMBDA=',E12.3)
      SSDIF=SSNEW-SSOLD
C              CONVERGENCE CRITERIA
      IF(NIT.EQ.ITMAX) GO TO 500
      NIT=NIT+1
      IF(SSDIF) 400,360,360
360   JTRIG=1
      IF(SSDIF.NE.0) GOTO 380
      JTR=JTR+1
      IF(JTR-3) 390,390,500
380   JTR=1
390   AMDA=AMDA*4.
      DO 370 J=1,NPAR
370   PAR(J)=PAR(J)-DX(J)
      CALL CAT(PAR)
      GO TO 10
400   AMDA=AMDA/SPEED
      SSOLD=SSNEW
      JTRIG=JTRIG+1
      IF(JTRIG-3) 10,10,450
450   IF(ABS(SSDIF/SSNEW)-TOL(2)) 500,500,480
480   IF(LGA) 500,500,10
500   CONTINUE
      SIGMA=SSNEW/FLOAT(NOB)
      CALL CI(A,XTX,D)
      DO 510 I=1,NPAR
      BNDUP(I)=PAR(I)+D(I)
510   BNDLW(I)=PAR(I)-D(I)
      RETURN
      END
C-----------------------------------------------------------------
C
C        ROUTINE FOR COMPUTING RESIDUAL A(T)
C
C-----------------------------------------------------------------
            SUBROUTINE CAT(PAR)
```

```
      DIMENSION PAR(1)
      COMMON/BLOK1/NAR,NMA,NOB,NPAR
      COMMON/BLOK4/X(1024),OBS(1024),AT(1024)
      IF(NPAR-NMA-NAR) 101,103,101
101   AVE=PAR(NPAR)
      DO 102 J=1,NOB
102   X(J)=OBS(J)-AVE
103   IF(NAR-NMA) 1,1,2
1     NMN=NMA
      GO TO 3
2     NMN=NAR
3     DO 15 I=1,NMN
      TEMP=0.
      IF(I-2) 15,5,5
5     II=I-1
      IF(II-NAR) 6,6,9
6     DO 8 J=1,II
      IJ=I-J
8     TEMP=TEMP-PAR(J)*X(IJ)
9     IF(II-NMA) 10,10,15
10    DO 12 J=1,II
      IJ=I-J
      JN=J+NAR
12    TEMP=TEMP+PAR(JN)*AT(IJ)
15    AT(I)=X(I)+TEMP
      NMN1=NMN+1
      DO 40 I=NMN1,NOB
      TEMP=0.
      DO 20 J=1,NAR
      IJ=I-J
20    TEMP=TEMP-PAR(J)*X(IJ)
      DO 30 J=1,NMA
      JN=J+NAR
      IJ=I-J
30    TEMP=TEMP+PAR(JN)*AT(IJ)
40    AT(I)=X(I)+TEMP
      RETURN
      END
C-----------------------------------------------------------------
C
C     SOLUTION OF THE SIMULTANEOUS EQUATION XB=Y
C
C-----------------------------------------------------------------
```

```
      SUBROUTINE EQN(X,M,B,Y)
      DIMENSION Y(40),B(40),X(40,40)
      CALL MATINV(X,M)
      DO 20 J=1,M
      B(J)=0.
      DO 20 K=1,M
20    B(J)=B(J)+Y(K)*X(J,K)
      RETURN
      END
C-----------------------------------------------------------------------
C                   MATRIX INVERSION ROUTINE
C
C-----------------------------------------------------------------------
      SUBROUTINE MATINV(A,N)
      DIMENSION A(40,40)
      NY=40
      DO 5 L=1,N
      X=A(L,L)
      A(L,L)=1.
      DO 2 J=1,N
2     A(L,J)=A(L,J)/X
      DO 5 K=1,N
      IF(K-L)3,5,3
3     X=A(K,L)
      A(K,L)=0.
      DO 4 J=1,N
4     A(K,J)=A(K,J)-X*A(L,J)
5     CONTINUE
      RETURN
      END
C-----------------------------------------------------------------------
C                ROUTINE FOR COMPUTING JACOBIAN
C-----------------------------------------------------------------------
      SUBROUTINE JAC(PAR,W,JCOL)
      DIMENSION PAR(40),THETA(20),W(1024)
      COMMON/BLOK1/NAR,NMA,NOB,NPAR
      COMMON/BLOK4/X(1024),OBS(1024),AT(1024)
      W(1)=0.
      NR=NAR+1
```

```
      DO 66 L=1,NMA
      LL=NAR+L
66    THETA(L)=PAR(LL)
      IF(JCOL-NMA-NAR)7,7,6
6     TQ=-1.
      DO 449 J=1,NAR
449   TQ=TQ+PAR(J)
      W(1)=TQ
7     DO 1 J=2,NAR
      IF(JCOL-NAR)22,22,2
22    M=J-JCOL
      TE=0.
      DO 4 K=1,NMA
      KK=J-K
      IF(KK)4,4,44
44    TE=TE+THETA(K)*W(KK)
4     CONTINUE
      W(J)=TE
      IF(M)1,1,62
62    W(J)=TE-X(M)
      GOTO 1
2     IF(JCOL-NMA-NAR)8,8,88
88    TE=0.
      DO 9 K=1,NMA
      KK=J-K
      IF(KK)9,9,99
99    TE=TE+THETA(K)*W(KK)
9     CONTINUE
      W(J)=TQ+TE
      GOTO 1
8     M=J-JCOL+NAR
      TE=0.
      DO 5 K=1,NMA
      KK=J-K
      IF(KK)5,5,55
55    TE=TE+THETA(K)*W(KK)
5     CONTINUE
      W(J)=TE
      IF(M)1,1,61
61    W(J)=AT(M)+TE
1     CONTINUE
      IF(JCOL-NAR)71,71,75
```

```
75         IF(JCOL-NMA-NAR)73,73,72
71         DO 11 J=NR,NOB
           M=J-JCOL
           TE=O.
           DO 84 K=1,NMA
           KK=J-K
84         TE=TE+THETA(K)*W(KK)
11         W(J)=TE-X(M)
           RETURN
72         DO 12 J=NR,NOB
           TE=O.
           DO 89 K=1,NMA
           KK=J-K
89         TE=TE+THETA(K)*W(KK)
12         W(J)=TQ+TE
           RETURN
73         DO 13 J=NR,NOB
           M=J-JCOL+NAR
           TE=O.
           DO 85 K=1,NMA
           KK=J-K
85         TE=TE+THETA(K)*W(KK)
13         W(J)=AT(M)+TE
           RETURN
           END
C-----------------------------------------------------------------

C     ROUTINE FOR COMPUTING CONFIDENCE INTERVAL

C-----------------------------------------------------------------
                SUBROUTINE CI(A,B,D)
           DOUBLE PRECISION GRD,SIGMA
           DIMENSION A(40,40),B(40,40)
           DIMENSION C(40,40),D(40)
           COMMON/BLOK1/NAR,NMA,NOB,NPAR
           COMMON/BLOK3/SIGMA,GRD(40),BNDLW(40),
     +     BNDUP(40),ITMAX,TOL(2),VAR
           DO 15 I=1,NPAR
           DO 15 J=1,NPAR
           C(I,J)=0.
           DO 10 K=1,NPAR
10         C(I,J)=C(I,J)+A(I,K)*B(K,J)
```

```
15          CONTINUE
C           D=C*AT^=A(XTX)AT^
            DO 20 I=1,NPAR
            D(I)=0.
            DO 21 J=1,NPAR
21          D(I)=D(I)+C(I,J)*A(I,J)
            SIG=SNGL(SIGMA)
20          D(I)=2.*SQRT(D(I)*SIG)
            RETURN
            END
C                   ********************
C                   *                  *
C                   *   20.   PMPSD    *
C                   *                  *
C                   ********************

            CHARACTER JWIN(6,4)
            DIMENSION XA(1024),XD(1024),XFR(513)
            DIMENSION SXX(513),WD(1024),ILAG(513)
            DIMENSION XREAL(1024),XIMAG(1024)
            REAL REXMN,IMXMN
            DATA JWIN(1,1),JWIN(1,2),JWIN(1,3),
     +            JWIN(1,4)/'R','E','C','T'/
            DATA JWIN(2,1),JWIN(2,2),JWIN(2,3),
     +            JWIN(2,4)/'H','A','M','G'/
            DATA JWIN(3,1),JWIN(3,2),JWIN(3,3),
     +            JWIN(3,4)/'B','H','A','M'/
            DATA JWIN(4,1),JWIN(4,2),JWIN(4,3),
     +            JWIN(4,4)/'H','A','N','N'/
            DATA JWIN(5,1),JWIN(5,2),JWIN(5,3),
     +            JWIN(5,4)/'T','R','I','A'/
            DATA JWIN(6,1),JWIN(6,2),JWIN(6,3),
     +            JWIN(6,4)/'P','A','R','Z'/
3           OPEN(1,FILE='   ')
            MAXM=1024
            LHM=MAXM/2+1
C           FILL LAG ARRAY FOR PRINTING
            DO 10 I=1,LHM
            ILAG(I)=I-1
10          CONTINUE
20          CONTINUE
```

```
      WRITE(*,999)
999   FORMAT(16H FFT LENGTH---M=)
      READ(*,*) M
      IF(M.GT.MAXM) WRITE(*,998)
998   FORMAT(27H M TOO LARGE--REENTER VALUE)
      IF(M.GT.MAXM)GOTO 20
      WRITE(*,996)
996   FORMAT(' WINDOW TYPE---1=RECTANGULAR,',
     *  ' 2=HAMMING',/' 3=GENERALIZED HAMMING,',
     *  ' 4=HANNING',/' 5=TRIANGULAR, 6=PARZEN')
      READ(*,*) IWIN
      WRITE(*,994)
994   FORMAT(19H WINDOW LENGTH---L=)
      READ(*,*) L
      WRITE(*,993)
993   FORMAT(' MAX. NO. OF DATA --- N=')
      READ(*,*) N
      WRITE(*,991)
991   FORMAT(' SAMPLING FREQ. IN HZ---FS=')
      READ(*,*) FS
C     NSECT = TOTAL NO. OF ANALYSIS SECTIONS
C        NP = TOTAL NO. OF SAMPLES USED
C        NP = N IF (N-L/2)/(L/2) = AN INTEGER
      MHLF1=M/2+1
      NSECT=(N-L/2)/(L/2)
      NP=NSECT*(L/2)+L/2
      WRITE(*,989)JWIN(IWIN,1),JWIN(IWIN,2),
     +   JWIN(IWIN,3),JWIN(IWIN,4)
989   FORMAT(13H WINDOW TYPE=,4A1)
      WRITE(*,988)M,NP,L,FS
988   FORMAT(' M=',I4,' NP=',I5,' L=',I4,
     1  ' SAMPLING FREQUENCY=',F10.4)
      IS=1
      READ(1,*) (XD(I),I=1,N)
C     CACULATE MEAN OF DATA
      XSUM=0.
      NS1=NSECT+1
      L1=L/2
      DO 40 K=1,NS1
      CALL GETD(XA,XD,L,IS)
      IS=IS-L
      DO 30 I=1,L1
```

```
      XSUM=XSUM+XA(I)
30    CONTINUE
      IS=IS+L1
40    CONTINUE
      XMEAN=XSUM/FLOAT(NP)
      REXMN=XMEAN
      IMXMN=XMEAN
      WRITE(*,986) XMEAN
986   FORMAT(7H XMEAN=,E14.5)
C     GENERATE WINDOW
      U=FLOAT(L)
      IF(IWIN.EQ.1)GOTO 60
      U=0.
      IF(IWIN.NE.3)GOTO 2000
      WRITE(*,2001)
2001  FORMAT(' SPECIFY ALFHA FOR GENERALIZED'.
     + ,' HAMMING WINDOW --- ALPHA = ')
      READ(*,*) ALPHA
2000  IW=IWIN-1
      TPI=8.*ATAN(1.)
      DO 50 I=1,L1
      FI=FLOAT(I-1)
      FL=FLOAT(L1-I)
      GOTO (221,222,223,224,225),IW
221   ALPHA=0.54
222   GOTO 226
223   ALPHA=.5
      LL=L
      GOTO 228
226   LL=L-1
228   BETA=1.-ALPHA
      WD(I)=ALPHA-BETA*COS(TPI*FI/FLOAT(LL))
      GOTO 227
224   WD(I)=1.-2.*FL/FLOAT(L)
      GOTO 227
225   WD(I)=1.-4.*(FL/FLOAT(L))**2
227   U=U+WD(I)*WD(I)
50    CONTINUE
      L11=L1+1
      DO 55 I=L11,L
      J=L-I+1
55    WD(I)=WD(J)
```

```
      U=2.*U
60    CONTINUE
C     LOOP TO ACCUMULATE SPECTRA 2 AT A TIME.
      IS=1
      DO 70 I=1,MHLF1
      SXX(I)=0.
70    CONTINUE
      NRD=L/2
      L2=L/2+1
      CALL GETD(XA(L2),XD,NRD,IS)
      IMN=L/2+1
      KMX=(NSECT+1)/2
      NSECTP=(NSECT+1)/2
      NRD=L
      DO 190 K=1,KMX
C     MOVE DOWN UPPER HALF OF XA BUFFER
      DO 80 I=1,L1
      J=L1+I
      XREAL(I)=XA(J)
      XIMAG(I)=0.
80    CONTINUE
      IF(K.NE.KMX.OR.NSECTP.EQ.NSECT)GOTO 100
      DO 90 I=IMN,NRD
      XA(I)=0.
90    CONTINUE
      NRD=L/2
100   CALL GETD(XA,XD,NRD,IS)
      DO 110 I=1,L1
      J=I+L1
      XREAL(J)=XA(I)-REXMN
      XIMAG(J)=XA(J)-IMXMN
      XREAL(I)=XREAL(I)-REXMN
      XIMAG(I)=XA(I)-IMXMN
110   CONTINUE
      IF(K.NE.KMX.OR.NSECTP.EQ.NSECT)GOTO 130
      DO 120 I=1,L
      XIMAG(I)=0.
120   CONTINUE
130   CONTINUE
      IF(IWIN.EQ.1)GOTO 150
      DO 140 I=1,L
      XREAL(I)=XREAL(I)*WD(I)
```

```
      XIMAG(I)=XIMAG(I)*WD(I)
140   CONTINUE
150   CONTINUE
      IF(L.EQ.M)GOTO 170
      LP1=L+1
      DO 160 I=LP1,M
      XREAL(I)=0.
      XIMAG(I)=0.
160   CONTINUE
170   CONTINUE
      CALL FFT(XREAL,XIMAG,M,0)
      DO 180 I=2,MHLF1
      J=M+2-I
      SXX(I)=SXX(I)+XREAL(I)*XREAL(I)+XIMAG(I)
     +   *XIMAG(I)
      SXX(I)=SXX(I)+XREAL(J)*XREAL(J)+XIMAG(J)
     +   *XIMAG(J)
180   CONTINUE
      SXX(1)=SXX(1)+XREAL(1)*XREAL(1)*2.
      SXX(1)=SXX(1)+XIMAG(1)*XIMAG(1)*2.
190   CONTINUE
C     NORMALIZE SPECTRAL ESTIMATE
C       AND OBTAIN CORRELATION FUNCTION
C    USING INVERSE FFT OF POWER SPECTRUM
      FNORM=2.*U*FLOAT(NSECT)
      SXX(1)=SXX(1)/FNORM
      XREAL(1)=SXX(1)
      XIMAG(1)=0.
      DO 200 I=2,MHLF1
      SXX(I)=SXX(I)/FNORM
      XREAL(I)=SXX(I)
      XIMAG(I)=0.0
      J=M+2-I
      XREAL(J)=XREAL(I)
      XIMAG(J)=XIMAG(I)
200   CONTINUE
      CALL FFT(XREAL,XIMAG,M,1)
      DO 210 I=1,MHLF1
      XA(I)=XREAL(I)
210   CONTINUE
C     CORRELATION ESTIMATE IS IN XA
C     COMPUTE LOG OF POWER SPECTRUM ESTIMATE
```

```
          XFS=FS/FLOAT(M)
          DO 220 I=1,MHLF1
          XFR(I)=FLOAT(I-1)*XFS
          TMP=ALOG10(SXX(I))
          SXX(I)=20.*TMP
220       CONTINUE
C         LOG POWER SPECTRUM(DB) IS IN ARRAY SXX
C         PLOT THE LOG POWER SPECTRUM
          WRITE(*,235)
235       FORMAT(17H PLOT PSD(1=YES)?)
          READ(*,*) IP
          IF(IP.NE.1)GOTO 9985
          CALL GRAPH(SXX,MHLF1)
9985      WRITE(*,985)
985       FORMAT(19H LOG POWER SPECTRUM)
          WRITE(*,984)
984       FORMAT(5X,'FREQ',7X,'DB',5X,'FREQ',7X,
     *    'DB',5X,'FREQ',7X,'DB',5X,'FREQ',7X,'DB')
          WRITE(*,983) (XFR(I),SXX(I),I=1,MHLF1)
983       FORMAT(4(F9.3,F9.4))
C         CORRELATION FUNCTION IS IN ARRAY XA
C         PLOT THE CORRELATION FUNCTION
          WRITE(*,245)
245       FORMAT(' PLOT CORRELATION FUNCTION ?',
     *    '(1 FOR YES)')
          READ(*,*) IC
          IF(IC.NE.1)GOTO 9982
          CALL GRAPH(XA,MHLF1)
9982      WRITE(*,982)
982       FORMAT(21H CORRELATION FUNCTION)
          WRITE(*,981)
981       FORMAT(1X,3HLAG,2X,4HCORR,5X,3HLAG,2X,
     *    4HCORR,5X,3HLAG,2X,4HCORR,5X,3HLAG,2X,
     *    4HCORR,5X,3HLAG,2X,4HCORR)
          WRITE(*,980) (ILAG(I),XA(I),I=1,MHLF1)
980       FORMAT(5(I4,E10.3))
          CLOSE(1)
          STOP
          END
```

```
C                    *****************
C                    *               *
C                    *  21.  CMPSD   *
C                    *               *
C                    *****************

        DIMENSION XA(512),XFR(257),W(512)
        CHARACTER*3 JWIN(8,4)
        DIMENSION XC(512),XS(512),ZC(257),ZS(257)
        DATA JWIN(1,1),JWIN(1,2),JWIN(1,3),
     +       JWIN(1,4)/'R','E','C','T'/
        DATA JWIN(2,1),JWIN(2,2),JWIN(2,3),
     +       JWIN(2,4)/'H','A','N','G'/
        DATA JWIN(3,1),JWIN(3,2),JWIN(3,3),
     +       JWIN(3,4)/'G','H','A','N'/
        DATA JWIN(4,1),JWIN(4,2),JWIN(4,3),
     +       JWIN(4,4)/'H','A','N','N'/
        DATA JWIN(5,1),JWIN(5,2),JWIN(5,3),
     +       JWIN(5,4)/'T','R','I','A'/
        DATA JWIN(6,1),JWIN(6,2),JWIN(6,3),
     +       JWIN(6,4)/'P','A','R','Z'/
        DATA JWIN(7,1),JWIN(7,2),JWIN(7,3),
     +       JWIN(7,4)/'K','A','I','S'/
        DATA JWIN(8,1),JWIN(8,2),JWIN(8,3),
     +       JWIN(8,4)/'3','-','C','O'/
50      WRITE(*,996)
996     FORMAT(' TOTAL NO. OF CORRELATION---N=')
        READ(*,*) N
        OPEN(1,FILE='      ')
        READ(1,*) (XA(I),I=1,N)
        CLOSE(1)
        WRITE(*,992)
992     FORMAT(31H SAMPLING FREQUENCY IN Hz---FS=)
        READ(*,*) FS
        WRITE(*,984)
984     FORMAT(' WINDOW TYPE ------ 1=RECTANGULAR,'
     1  ,' 2=HAMMING,'/,2X,'3=GENERALIZED HAMMING,'
     +  ,' 4=HANNING, 5=TRIANGUALAR, 6=PARZEN,'/
     1  ,2X,'7=MODIFIED KAISER, 8=3-COEFFICIENT')
        READ(*,*) IWIN
        WRITE(*,982)
```

```
982     FORMAT(' NO. OF CORRELATION VALUES USED',
     +   '---L=')
        READ(*,*) L
        WRITE(*,980)
980     FORMAT(16H FFT SIZE---NFFT)
        READ(*,*) NFFT
        NHLF1=NFFT/2+1
        WRITE(*,979)JWIN(IWIN,1),JWIN(IWIN,2),
     *  JWIN(IWIN,3),JWIN(IWIN,4),L,NFFT
979     FORMAT(' WINDOW TYPE=',4A1,'  NO. OF '
     *  ,'WINDOW VALUES=',I4,' FFT SIZE=',I4)
C       WINDOW CORRELATION FUNCTION
        PI=4.0*ATAN(1.0)
        IF(IWIN.NE.7) GOTO 229
        WRITE(*,2221)
2221    FORMAT(2X,'SPECIFY BETA FOR MODIFIED '
     +  ,'KAISER WINDOW ( 0<=BETA<=10 )')
 231    READ(*,*) BETA
        IF(IWIN.EQ.8) GOTO 215
        CALL MODKAI(L,W,BETA)
 229    IF(IWIN.NE.8) GOTO 227
        WRITE(*,2220)
2220    FORMAT(2X,'SPECIFY BETA FOR 3-COEF WINDOW'
     +  ,' (0<=BETA<=0.045)')
        GOTO 231
227     IF( IWIN.NE.3) GOTO 215
        WRITE(*,2222)
2222    FORMAT(2X,'SPECIFY ALPHA FOR GENERALIZED'
     +  ,' HAMMING WINDOW'/)
        READ(*,*) ALPHA
215     DO 230 I=2,L
        GOTO (220,221,222,223,224,225,228,232),
     +   IWIN
221     ALPHA=0.54
        GOTO 222
223     ALPHA=0.5
        LL=L
        GOTO 226
222     LL=L-1
226     BETA=1.-ALPHA
        XA(I)=XA(I)*(ALPHA+BETA*COS(PI*FLOAT(I-1)
     +          /FLOAT(LL)))
```

```
      GOTO 220
224   XA(I)=XA(I)*(1.0-FLOAT(I-1)/FLOAT(L))
      GOTO 220
225   XA(I)=XA(I)*(1.-(FLOAT(I-1)/
     + FLOAT(L))**2)
      GOTO 220
228   XA(I)=XA(I)*W(I)
      GOTO 220
232   T=FLOAT(I-1)/FLOAT(L-1)
      WI=.5*COS(PI*T)+2.*BETA*COS(2.*PI*T)
      WI=.5+WI-2.*BETA
      XA(I)=XA(I)*WI
220   IF(MODE.EQ.1.OR.MODE.EQ.3)GOTO 230
      J=NFFT+2-I
      XA(J)=XA(I)
230   CONTINUE
      NLAST=NFFT+1-L
      IF(MODE.EQ.1.OR.MODE.EQ.3)NLAST=NFFT
      L1=L+1
      DO 240 I=L1,NLAST
      XA(I)=0.
240   CONTINUE
      DO 250 I=1,NFFT
      XC(I)=XA(I)
      XS(I)=0.
250   CONTINUE
      CALL FFT(XC,XS,NFFT,0)
      XFS=FS/FLOAT(NFFT)
      NHF=NFFT/2
      NHF1=NHF+1
      DO 260 I=1,NHF1
      XFR(I)=FLOAT(I-1)*XFS
      T=ALOG10(SQRT(XC(I)*XC(I)+XS(I)*XS(I)))
      XA(I)=20.*T
260   CONTINUE
C     LOG POWER SPECTRUM (DB) IS IN XA
C     PRINT AND PLOT LOG POWER SPECTRUM
      WRITE(*,265)
265   FORMAT(' PLOT POWER SPECTRUM?(1=YES,0=NO)')
      READ(*,*) ID
      IF(ID.EQ.0) GOTO 9978
      CALL GRAPH(XA,NHLF1)
9978  WRITE(*,978)
```

```
978       FORMAT(19H LOG POWER SPECTRUM)
          WRITE(*,977)
977       FORMAT(5X,4HFREQ,5X,2HDB,7X,4HFREQ,5X,2
     *    HDB,7X,4HFREQ,5X,2HDB,7X,4HFREQ,5X,2HDB)
          WRITE(*,976) (XFR(I),XA(I),I=1,NHLF1)
976       FORMAT(4(F9.3,F9.4))
          STOP
          END
C*********************************************************
          SUBROUTINE MODKAI(L,W,BETA)
          DIMENSION W(1)
          REAL IO
          BESL=IO(BETA)-1.
          XIN=FLOAT(L-1)*FLOAT(L-1)
          DO 10 I=1,L
          XI=(I-1)**2
          W(I)=IO(BETA*SQRT(1.-XI/XIN))
          W(I)=(W(I)-1.)/BESL
10        CONTINUE
          RETURN
          END
C*********************************************************
          REAL FUNCTION IO(X)
          Y=X/2.
          T=1.E-08
          E=1.
          DE=1.
          DO 10 I=1,25
             XI=I
             DE=DE*Y/XI
             SDE=DE*DE
             E=E+SDE
             IF(E*T-SDE) 10,10,20
10        CONTINUE
20        IO=E
          RETURN
          END


C                    ******************
C                    *                *
C                    *  22.  MEBURG  *
C                    *                *
C                    ******************
```

```
      DIMENSION X(1024),FPE(50),G(50)
      DIMENSION A(50),Y(1024),PS(500)
      OPEN(1,FILE='  ')
      PI2=8.*ATAN(1.)
      WRITE(*,9995)
9995  FORMAT(' NO. OF DATA --- N = ?(<=1024)')
      READ(*,*)N
      READ(1,*) (X(I),I=1,N)
1111  WRITE(*,9993)
9993  FORMAT(' SUBTRACTS MEAN ---IM = ?',
     +'(YES=1,NO=0)')
      READ(*,*)IM
      IF(IM.EQ.0)GOTO 14
      AVE=0.
      DO 13 I=1,N
      AVE=AVE+X(I)
13    CONTINUE
      AVE=AVE/FLOAT(N)
      DO 15 I=1,N
15    Y(I)=X(I)-AVE
      GO TO 16
14    DO 12 I=1,N
12    Y(I)=X(I)
16    WRITE(*,9991)
9991  FORMAT(' HIGHEST ORDER---MLAST=(<40)')
      READ(*,*)MLAST
      WRITE(*,9989)
9989  FORMAT(' FORMULA FOR ORDER SELECTION',
     +' --- IC = 1 OR 2')
      READ(*,*)IC
9982  FORMAT(3X,'N = ?',I5,5X,'IM = ?',I3,5X,
     +'IC = ?',I2)
      LG=MLAST+1
      LEXT=LG+1
      CALL BURG(N,Y,LG,G,A,LEXT,FPE,PM,IC,IM)
      FPEMIN=1E15
      DO 30 I=2,LG
      IF(FPE(I).GT.FPEMIN)GOTO 30
      FPEMIN=FPE(I)
      NAROPT=I-1
30    CONTINUE
      WRITE(*,40)NAROPT
```

```
40        FORMAT(1X,' OPTIMUM ORDER =',I4)
          WRITE(*,9992)
9992      FORMAT(' CALCULATES PSD?(YES=1,NO=0)')
          READ(*,*)IS
          IF(IS.EQ.0)GOTO 2222
          WRITE(*,9984)
9984      FORMAT(' STARTING VALUE OF FREQUENCY',
     +     ' --- FO = ?')
          READ(*,*) FO
          WRITE(*,9990)
9990      FORMAT(' FREQUENCY INTERVAL---DF=?')
          READ(*,*)DF
9985      FORMAT(F8.5)
          WRITE(*,9988)
9988      FORMAT(' NO. OF PSD POINTS TO BE PRINTED',
     +     ' --- NF = ?')
          READ(*,*)NF
          NF1=NF+1
          F=FO
          DO 50 I=1,NF1
          SUMR=1.
          SUMI=0.
          DO 60 J=1,MLAST
          SUMR=SUMR+G(J+1)*COS(PI2*F*J)
          SUMI=SUMI-G(J+1)*SIN(PI2*F*J)
60        CONTINUE
          PSD=PM/(SUMR*SUMR+SUMI*SUMI)
          PS(I)=PSD
          WRITE(*,9987)F,PSD
          F=F+DF
50        CONTINUE
9987      FORMAT(6H FREQ=,E25.8,8H    PSD=,E25.8)
          WRITE(*,9983)
9983      FORMAT('     FREQ      PSD')
          CALL GRAPH(PS,NF)
2222      WRITE(*,9986)
9986      FORMAT(//,' MORE FIT ? (1=YES,0=NO)=')
          READ(*,*)IMD
          IF(IMD.EQ.1)GOTO 1111
          CLOSE(1)
          STOP
          END
```

```
C************************************
        SUBROUTINE OUTPUT(NN,G,FPE,PM)
        DIMENSION GG(50),FPE(50),G(50)
        NAR=NN-1
        WRITE(*,100)NAR,FPE(NN),PM
100     FORMAT(1X,'AR ORDER =',I2,5X,'FPE=',
     +    F15.9,5X,'P =',F15.9)
        DO 44 I=2,NN
        GG(I)=-G(I)
44      CONTINUE
        WRITE(*,201)  (GG(I),I=2,NN)
201     FORMAT(5H PHI=,F15.9)
        RETURN
        END
C*************************************
        SUBROUTINE BURG(M,F,LG,G,A,LEXT,FPE,PM,
     +   IC,IM)
        REAL F(M),FPE(M),G(M),A(LEXT),PEF(1200)
        REAL PER(1200),H(1200)
        SUM=0.
        DO 13 I=1,M
        T1=F(I)
13      SUM=SUM+T1*T1
        A(1)=SUM/FLOAT(M)
        PM=A(1)
        DM=A(1)
        IF(IM.EQ.1)GOTO 40
        FPE(1)=FLOAT(M+1)/FLOAT(M-1)*PM
        IF(IC.EQ.2)FPE(1)=FLOAT(M+2)/FLOAT(M-2)
     +  *PM
        GOTO 30
40      CONTINUE
        FPE(1)=FLOAT(M+2)/FLOAT(M-2)*PM
        IF(IC.EQ.2)FPE(1)=FLOAT(M+4)/FLOAT(M-4)
     +  *PM
30      FTEMP=FPE(1)
        FPE(1)=0.
        DO 100 NN=2,LG
        N=NN-2
        IF(N.NE.0)GO TO 11
        DO 12 J=1,M
        PEF(J)=0.
```

```
12        PER(J)=0.
11        SN=0.
          SD=0.
          JJ=M-N-1
          DO 200 J=1,JJ
          JN1=J+N+1
          T1=F(JN1)
          T2=F(J)
          SN=SN-2.*(T1+PEF(J))*(T2+PER(J))
200       SD=SD+(T1+PEF(J))**2+(T2+PER(J))**2
          G(NN)=SN/SD
          IF(ABS(G(NN)).GT.1.)GO TO 101
          IF(N.EQ.0)GO TO 3
          DO 4 J=1,N
          K=N-J+2
4         H(J+1)=G(J+1)+G(NN)*G(K)
          DO 6 J=1,N
6         G(J+1)=H(J+1)
          JJ=JJ-1
3         DO 10 J=1,JJ
          JNN1=J+NN-1
          PER(J)=PER(J)+G(NN)*PEF(J)+G(NN)*F(JNN1)
10        PEF(J)=PEF(J+1)+G(NN)*PER(J+1)+G(NN)*
     +    F(J+1)
          SUM=0.
          DO 14 J=2,NN
          NNJ1=NN+1-J
14        SUM=SUM-A(NNJ1)*G(J)
          A(NN)=SUM
          DM=(1.-G(NN)**2)*DM
          PM=DM
          IF(NN.EQ.M) GO TO 88
          IF(IC.EQ.2) GOTO 50
          MPNN=M+NN
          MMNN=M-NN
          IF(IM.EQ.0)GOTO 20
          MPNN=MPNN+1
          MMNN=MMNN-1
          GOTO 20
50        MPNN=M+NN*2
          MMNN=M-NN*2
          IF(IM.EQ.0) GOTO 20
```

```
      MPNN=MPNN+2
      MMNN=MMNN-2
      GOTO 20
20    FPE(NN)=FLOAT(MPNN)/FLOAT(MMNN)*PM
      FPE(NN)=FPE(NN)/FTEMP
      FPE(NN)=ALOG10(FPE(NN))
88    CONTINUE
      CALL OUTPUT(NN,G,FPE,PM)
100   CONTINUE
      G(1)=1.
      LG1=LG+1
      DO 16 J=LG1,LEXT
      SUM=0.
      DO 17 I=2,LG
      JI1=J+1-I
17    SUM=SUM-A(JI1)*G(I)
16    A(J)=SUM
99    CONTINUE
      RETURN
101   LG=NN-2
      WRITE(*,333) LG
333   FORMAT(1X,'NUMERICAL INSTABILITY OCCURED;'
     * HIGHEST ORDER FITTED =',I3)
      RETURN
      END


C                  ********************
C                  *                  *
C                  *   23.  INVERF    *
C                  *                  *
C                  ********************


      DIMENSION F(50),C(50),RI(50),RIO(50),
     * WP(50),RL(50),W(2000),LL(50)
      OPEN(1,FILE=' ')
      WRITE(*,101)
101   FORMAT(2X' AR- AND MA- ORDER OF THE',
     * ' MODEL  IP=?  IQ=?')
      READ(*,*)IP,IQ
      IF(IP.EQ.0)GOTO 20
```

```
      WRITE(*,105)
105   FORMAT(2X,'AR- PARAMETERS OF THE'
     *,' MODEL    :')
      READ(*,*)(F(I),I=1,IP)
20    IF(IQ.EQ.0)GOTO 21
      WRITE(*,106)
106   FORMAT(2X,'MA- PARAMETERS OF THE'
     *,' MODEL   :')
      READ(*,*)(C(I),I=1,IQ)
21    WRITE(*,102)
102   FORMAT(2X,'START POINT AND END POINT '
     *,'OF FORECAST  K=?  KK=?')
      READ(*,*)K,KK
      READ(1,*)(W(I),I=1,KK)
      WRITE(*,103)
103   FORMAT(2X,'THE STEP OF FORECAST  L=?')
      READ(*,*)L
      WRITE(*,109)
109   FORMAT(2X,'PRECISION OF FORECAST'
     *,'   EPS=?')
      READ(*,*)EPS
      L1=L+1
      DO 22 I=1,L1
22    LL(I)=I-1
      WRITE(*,110)(LL(I),I=1,L1)
110   FORMAT(1X,'FORECAST',10X,'FORECAST'
     *,' STEP'/,2X,'TIME',10I10,/)
      MPQL=MAX0(IP,IQ,L)
      DO 10 K1=K,KK
      CALL INVERF(K1,L,MPQL,IP,IQ,F,C,
     *RI,RIO,RL,W,WK,EPS)
10    CONTINUE
      CLOSE(1)
      STOP
      END
C*****************************************************
      SUBROUTINE INVERF(K,L,MPQL,IP,IQ,
     *F,C,RI,RIO,RL,W,WK,EPS)
      REAL RI(MPQL),RIO(MPQL),RL(MPQL),
     *F(MPQL),C(MPQL),W(K),WK(MPQL)
      DO 10 I=1,MPQL
      WK(I)=0.
```

```
10          RI(I)=0.
            RI(1)=-1.
            II=1
            JJ=1
            RIW=O.
            L1=L-1
200         RO=0.
            IF(L.EQ.0)GOTO 16
            IF(II.LE.IP)RO=F(II)
            DO 12 I1=1,MPOL
            I=MPOL-I1+1
            IF(I.LE.IQ)RO=RO+C(I)*RI(I)
            IF(I.NE.1)RI(I)=RI(I-1)
12          CONTINUE
            RI(1)=RO
            II1=L-II+1
            IF(II1)14,14,18
18          RIO(II1)=RO
            IF(II.NE.L)GOTO 13
14          CONTINUE
            DO 15 I=1,L
            I2=L-I+1
            RILO=RI(I2)
            I1=I-1
            IF(I1.EQ.0)GOTO 15
            DO 17 J=1,I1
            J1=L-J+1
            J2=I1-J+1
17          RILO=RILO+RIO(J1)*RL(J2)
15          RL(I)=RILO
            KJ=K-JJ+1
            WKJ=O.
            IF(KJ.GT.O)WKJ=W(KJ)
            RIWO=-1.E8
            DO 19 I=1,MPOL
            RIW=RL(I)*WKJ
            RA=ABS(RIW)
            WK(I)=WK(I)+RIW
            IF(RIWO.LE.RA)RIWO=RA
19          CONTINUE
16          IF(ABS(RIWO).LE.EPS)GOTO 201
            JJ=JJ+1
```

```
13        II=II+1
          GOTO 200
201       WRITE(*,105)K,W(K),(WK(I),I=1,I)
105       FORMAT(1X,I4,5X,10G10.4)
          RETURN
          END
```

```
C                   ******************
C                   *                *
C                   *  24.  VECTF    *
C                   *                *
C                   ******************

          DIMENSION F(50),C(50),WW(50),
     *    W(2000),LL(50),G(50)
          OPEN(1,FILE=' ')
          OPEN(2,FILE=' ')
          WRITE(*,101)
101       FORMAT(2X,' AR- AND MA- ORDER OF THE',
     *    ' MODEL  IP=?  IQ=?')
          READ(*,*)IP,IQ
          IF(IP.EQ.0)GOTO 20
          WRITE(*,105)
105       FORMAT(2X,'AR- PARAMETERS OF THE'
     *    ,' MODEL    :')
          READ(*,*)(F(I),I=1,IP)
20        IF(IQ.EQ.0)GOTO 21
          WRITE(*,106)
106       FORMAT(2X,'MA- PARAMETERS OF THE'
     *    ,' MODEL   :')
          READ(*,*)(C(I),I=1,IQ)
21        WRITE(*,102)
102       FORMAT(2X,'START POINT AND END POINT'
     *    ,'OF FORECAST  K=?  KK=?')
          READ(*,*)K,KK
          READ(1,*)(W(I),I=1,KK)
          WRITE(*,103)
```

```
103       FORMAT(2X,'THE STEP OF FORECAST  L=?')
          READ(*,*)L
          READ(2,*)(WW(I),I=1,L)
          L1=L+1
          DO 22 I=1,L1
22        LL(I)=I-1
          WRITE(*,110)(LL(I),I=1,L1)
110       FORMAT(1X,'FORECAST',10X,'FORECAST'
     *    ,' STEP'/,2X,'TIME',10I10,/)
          MPQL=MAXO(IP,IQ,L)
          MPQL=MAXO(IP,IQ,L)
          DO 1 I1=1,MPQL
          I=MPQL-I1+1
          G(I)=0.
          KI=K-I1+1
1         WW(I)=W(KI)
          G(1)=1.
          CALL GREENF(IP,IQ,MPQL,L,F,C,G)
          DO 3 K1=K,KK
          WK1=W(K1)
          CALL VECTF(K1,L,MPQL,IP,IQ,F,C,G,
     *    WK,WW,WK1)
3         CONTINUE
          CLOSE(1)
          STOP
          END
C*******************************************************
          SUBROUTINE VECTF(K,L,MPQL,IP,IQ,
     *    F,C,G,WK,WW,WW1)
          REAL F(MPQL),C(MPQL),G(MPQL),
     *    WK(MPQL),WW(MPQL)
          WWK=WW1-WK(1)
          IQ1=IQ+1
          CALL PRED(IQ1,L,K,MPQL,IP,IQ,
     *    F,C,WK,WW,WJ)
3         DO 30 I=1,MPQL
          I1=MPQL-I+1
          IF(I.GT.IQ)GOTO 15
          WO=G(I1)*WWK
          IF(I.EQ.IQ)GOTO 11
          W1=WK(I+1)
          GOTO 12
```

```
 11     W1=WJ
 12     WO=WO+W1
        GOTO 30
 15     CALL PRED(I,L,K,MPOL,IP,IQ,
     *  F,C,WK,WW,WO)
 30     WK(I)=WO
        MPOL1=MPOL-1
        DO 17 J=1,MPOL1
 17     WW(J)=WW(J+1)
        WW(MPOL)=WW1
        WRITE(*,101)K,WW1,(WK(I),I=1,L)
 101    FORMAT(1X,I4,5X,10G10.4)
        RETURN
        END
C*********************************************************
        SUBROUTINE PRED(J,L,K,MPOL,IP,IQ,
     *  F,C,WK,WW,WJ)
        REAL F(MPOL),C(MPOL),WK(MPOL),WW(MPOL)
        WJ=0.
        IF(IP.EQ.0)GOTO 20
        DO 21 I=1,IP
        IO=J-I
        FI=F(I)
        IF(IO)10,11,11
 11     WK1=WK(IO)
        GOTO 24
 10     K1=MPOL+IO
        WK1=WW(K1)
 24     WJ=WJ+FI*WK1
 21     CONTINUE
 20     RETURN
        END

  C                 ******************
  C                 *                *
  C                 *   25.  MVAR    *
  C                 *                *
  C                 ******************

        REAL Z(600,8),R(30,8,8),RO(8,8),E(8),
```

```
     *    A(30,8,8),HA(8,8),HB(8,8),H(8,8),
     *    B(30,8,8),D(8),FPE(8),EE(8)
          DIMENSION FPEO(8),MO(8)
          OPEN(1,FILE=' ')
          WRITE(*,101)
101       FORMAT(2X,'N=?   K=?   M=?')
          READ(*,*)N,K,M
          READ(1,*)((Z(I,J),I=1,N),J=1,K)
          CALL MULCOV(N,K,M,Z,RO,R,E,EE)
          WRITE(*,102)
102       FORMAT(3X,'MAX-DIM',3X,'DIM',4X,
     *    'AR-',5X,'FPE',6X,'MUL.AR-'
     *    ,/2X,18X,'ORDER',
     *    13X,'PARAMETERS')
          DO 10 L1=1,K
          L=K-L1+1
          CALL MVAR(M,K,L,N,Z,R,RO,A,B,HA,HB,
     *    H,D,E,EE,FPE,FPEO,MO,ISW)
10        CONTINUE
          CLOSE(1)
          STOP
          END
C*************************************************************
          SUBROUTINE MULCOV(N,K,M,Z,RO,R,D,EE)
          REAL Z(600,8),RO(K,K),R(M,K,K),D(K),EE(K)
          TN=FLOAT(N)
          M1=M+1
          DO 16 I=1,K
          E=0.
          DO 17 J=1,N
17        E=E+Z(J,I)
          E=E/TN
16        EE(I)=E
          DO 10 J1=1,M1
          J=J1-1
          DO 12 I1=1,K
          DO 12 I2=1,K
          T=0.
          DO 13 IT=J1,N
          ITJ=IT-J
13        T=T+(Z(IT,I1)-EE(I1))*(Z(ITJ,I2)-EE(I2))
          T=T/TN
```

```
      IF(J)14,14,15
14    RO(I1,I2)=T
      T0-T
      GOTO 12
15    R(J,I1,I2)=T
12    CONTINUE
10    CONTINUE
      DO 19 I1=1,K
      RI11=RO(I1,I1)
      DO 19 I2=1,K
      R12=SQRT(RI11)*SQRT(RO(I2,I2))
      DO 19 J=1,M
19    R(J,I1,I2)=R(J,I1,I2)/R12
      DO 20 I1=1,K
      I11=I1+1
      R1=SQRT(RO(I1,I1))
      D(I1)=R1
      IF(K.EQ.1)GOTO 20
      DO 22 I2=I11,K
      T=RO(I1,I2)/(R1*SQRT(RO(I2,I2)))
      RO(I1,I2)=T
22    RO(I2,I1)=T
20    RO(I1,I1)=1.
      WRITE(*,100)((RO(I1,I2),I1=1,K),I2=1,K)
      DO 21 J=1,M
21    WRITE(*,100)((R(J,I1,I2),I1=1,K),I2=1,K)
100   FORMAT(16F5.2)
      WRITE(*,102)D
      WRITE(*,103)EE
102   FORMAT(//,5X,'R.M.S.       :',10F10.3,//)
103   FORMAT(5X,'MEAN          : ,10F10.3,//)
      RETURN
      END
C*********************************************************
      SUBROUTINE ABP(N,M,IP,K,L,R,RO,B,A,HB,
     *  HA,H,D,E,FPE,ISW)
      REAL RO(K,K),R(M,K,K),A(M,L,L),B(M,L,L)
     *  ,HB(L,L),HA(L,L),H(L,L),D(L),E(L),FPE(L)
      TN=FLOAT(N)
      KKK=0
      IP1=IP+1
66    DO 20 I1=1,L
```

```
      DO 20 I2=1,L
      T1=0.
      T2=0.
      IF(IP.EQ.0)GOTO 24
      DO 22 J=1,IP
      JP=IP+1-J
      DO 22 I3=1,L
      IF(KKK.EQ.0)T2=T2+A(J,I1,I3)*R(JP,I3,I2)
      IF(I2.GT.I1)GOTO 22
      IF(KKK)11,11,10
10    TT=B(J,I1,I3)*R(J,I3,I2)
      GOTO 23
11    TT=A(J,I1,I3)*R(J,I2,I3)
23    T1=T1+TT
22    CONTINUE
24    IF(IP.EQ.M)GOTO 15
      IF(KKK)13,13,14
13    HA(I1,I2)=R(IP1,I1,I2)-T2
      GOTO 15
14    IF(I2.GT.I1)GOTO 20
      T=HA(I1,I2)
      HA(I1,I2)=HA(I2,I1)
      HA(I2,I1)=T
15    IF(I2.GT.I1)GOTO 20
      T1=RO(I1,I2)-T1
      HB(I1,I2)=T1
      IF(I2.NE.I1)HB(I2,I1)=T1
20    CONTINUE
      CALL LDLF(L,L,HB,D,HA,H,ISW)
      IF(ISW.EQ.1)GOTO 44
      IF(KKK)50,50,51
50    DETO=1.
      TNLP=FLOAT(N+L*IP)/FLOAT(N-L*IP)
      DO 25 I1=1,L
      EI1=E(I1)
      DETO=DETO/D(I1)*EI1*EI1*TNLP
      FPE(I1)=DETO
      IF(IP.EQ.M)GOTO 44
      DO 25 I2=1,L
      B(IP1,I2,I1)=H(I1,I2)
25    CONTINUE
      KKK=1
```

```
      GOTO 55
51    DO 18 I1=1,L
      EI1=E(I1)
      DO 18 I2=1,I1
      EI2=E(I2)
      A(IP1,I2,I1)=H(I1,I2)
      IF(I1.EQ.I2)GOTO 18
      A(IP1,I1,I2)=H(I2,I1)
18    CONTINUE
44    RETURN
      END
C*********************************************************
      SUBROUTINE ABJ(M,IP,L,J,A,B,HA,HB)
      REAL A(M,L,L),B(M,L,L),HA(L,L),HB(L,L)
      JP=IP+1-J
      IP1=IP+1
      DO 10 I1=1,L
      DO 10 I2=1,L
      HA(I1,I2)=A(J,I1,I2)
10    HB(I1,I2)=B(JP,I1,I2)
      DO 11 I1=1,L
      DO 11 I2=1,L
      T1=0.
      T2=0.
      DO 12 I3=1,L
      T1=T1+A(IP1,I1,I3)*HB(I3,I2)
12    T2=T2+B(IP1,I1,I3)*HA(I3,I2)
      A(J,I1,I2)=HA(I1,I2)-T1
11    B(JP,I1,I2)=HB(I1,I2)-T2
      RETURN
      END
C*********************************************************
      SUBROUTINE ABPJ(N,M,K,L,IP,R,RO,A,B,HA,
     * HB,H,D,E,FPE,ISW)
      REAL RO(K,K),R(M,K,K),A(M,L,L),B(M,L,L),
     * E(L),HA(L,L),HB(L,L),H(L,L),D(L),FPE(L)
      CALL ABP(N,M,IP,K,L,R,RO,B,A,HB,HA,H,D,
     * E,FPE,ISW)
      IF(ISW.EQ.1)GOTO 44
      IF(IP.EQ.M1)GOTO 44
      IF(IP.EQ.0)GOTO 44
      I=IP/2
```

```
      IF(I.EQ.0)GOTO 5
      DO 23 J=1,I
      CALL ABJ(M,IP,L,J,A,B,HA,HB)
      JP=IP+1 -J
23    CALL ABJ(M,IP,L,JP,A,B,HA,HB)
5     IF(I*2-IP.EQ.0)GOTO 44
      J=I+1
      CALL ABJ(M,IP,L,J,A,B,HA,HB)
44    RETURN
      END
C****************************************************
      SUBROUTINE MVAR(M,K,L,N,Z,R,RO,A,B,HA,
     *  HB,H,D,E,EE,FPE,FPEO,MO,ISW)
      REAL Z(600,8),RO(K,K),R(M,K,K),A(M,L,L),
     *  B(M,L,L),E(L),HA(L,L),HB(L,L),H(L,L),
     *  D(L),FPE(L),FPEO(L),EE(L)
      DIMENSION MO(L)
      TN=FLOAT(N)
      DO 33 I=1,L
33    FPEO(I)=1.E10
      M1=M+1
      DO 21 IP1=1,M1
      IP=IP1-1
      CALL ABPJ(N,M,K,L,IP,R,RO,A,B,HA,HB,H,
     *  D,E,FPE,ISW)
      IF(ISW.EQ.1)GOTO 44
      IF(IP1.EQ.M1)GOTO 25
      DO 34 I=1,IP1
      DO 37 I1=1,L
      DO 37 I2=1,I1
      HA(I1,I2)=A(I,I1,I2)*E(I1)/E(I2)
      IF(I1.EQ.I2)GOTO 37
      HA(I2,I1)=A(I,I2,I1)*E(I2)/E(I1)
37    CONTINUE
34    CONTINUE
25    DO 21 I=1,L
      IF(FPEO(I).LE.FPE(I))GOTO 21
      FPEO(I)=FPE(I)
      MO(I)=IP
21    CONTINUE
      DO 10 I=1,L
      MI=MO(I)
```

```
      FPEOI=FPEO(I)
      WRITE(*,105)L,I,MI,FPEOI
105   FORMAT(2(5X,I2),7X,I2,4X,G8.3)
      DO 11 J1=1,MI
      J=J1-1
11    CALL ABPJ(N,M,K,L,J,R,RO,A,B,HA,HB,
     * H,D,E,FPE,ISW)
      IF(ISW.EQ.1)GOTO 44
      DO 16 I1=1,I
16    D(I1)=EE(I1)
      DO 15 J2=1,MI
      DO 15 I1=1,I
      DO 15 I2=1,I
15    D(I1)=D(I1)-A(J2,I1,I2)*E(I1)/E(I2)*EE(I2)
      WRITE(*,109)(D(I1),I1=1,I)
      DO 12 J2=1,MI
      DO 35 I1=1,I
      DO 35 I2=1,I1
      HA(I1,I2)=A(J2,I1,I2)*E(I1)/E(I2)
      IF(I1.EQ.I2)GOTO 35
      HA(I2,I1)=A(J2,I2,I1)*E(I2)/E(I1)
35.   CONTINUE
      WRITE(*,109)((HA(I1,I2),I1=1,I),I2=1,I)
12    CONTINUE
10    CONTINUE
109   FORMAT(33X,9F7.3)
44    RETURN
      END
C*************************************************************
      SUBROUTINE LDLT(N,M,A,P,B,X,ISW)
      REAL A(N,N),P(N),B(N,M),X(N,M)
      DO 5 I=1,N
      DO 5 J=1,I
      W=A(J,I)
      IF(I-J)10,20,10
20    IF(I.EQ.1)GOTO 30
      IM1=I-1
      DO 25 K=1,IM1
      Y=A(I,K)
      Z=Y*P(K)
      A(I,K)=Z
25    W=W-Y*Z
```

```
30        IF(W.NE.0.)GOTO 35
          ISW=1
          WRITE(*,101)ISW
101       FORMAT(2X,'ISW=',I1)
          RETURN
35        P(I)=1./W
          GOTO 5
10        IF(J.EQ.1)GOTO 45
          JM1=J-1
          DO 40 K=1,JM1
40        W=W-A(I,K)*A(J,K)
45        A(I,J)=W
5         CONTINUE
          DO 100 J=1,M
          DO 60 I=1,N
          Y=B(I,J)
          IF(I.EQ.1)GOTO 60
          IM1=I-1
          DO 65 K=1,IM1
65        Y=Y-A(I,K)*X(K,J)
60        X(I,J)=Y
          DO 100 L=1,N
          I=N-L+1
          Y=X(I,J)*P(I)
          IF(I.EQ.N)GOTO 100
          IP1=I+1
          DO 70 K=IP1,N
70        Y=Y-A(K,I)*X(K,J)
100       X(I,J)=Y
          ISW=0
          RETURN
          END
C                   *******************
C                   *                 *
C                   *   26.   SMLR    *
C                   *                 *
C                   *******************

          REAL S(75,75),SS(75),E(75),F(75),
     *    FF(75),X(1000,5)
          DIMENSION L(1000),LL1(75),LL2(75)
```

```
      COMMON /COM1/N,IS,X
      COMMON /COM2/M,KK,L
      OPEN(1,FILE=' ')
      WRITE(*,101)
      READ(*,*)N
      WRITE(*,102)
      READ(*,*)IS
      WRITE(*,103)
      READ(*,*)M
      WRITE(*,104)
      READ(*,*)KK
      READ(1,*)((X(I,J),I=1,N),J=1,IS)
      MRE=M*IS+1
      N1=M+1
      N2=N
      CALL RECOV(N1,N2,MRE,S,E,SS,A)
      CALL SELECT(MRE,S,SS,E,F,FF,BICO,
     *        II,ISW)
      IF(ISW.EQ.1)GOTO 999
      K=1
      DO 51 I=1,II
      LL=L(I)
      L1=LL/M
      L2=LL-L1*M
      IF(L2)3,3,4
3     L2=M
      GOTO 50
4     L1=L1+1
50    LL1(K)=L1
      LL2(K)=L2
51    K=K+1
      II1=II+1
      WRITE(*,105)
      WRITE(*,106)KK,F(II1)
      DO 6 I=1,II
6     WRITE(*,107)F(I),LL2(I),LL1(I)
101   FORMAT(2X,'NUMBER OF DATA IN'
     *  ,' EACH DIMENSION  N=?')
102   FORMAT(2X,'DIMENSION OF OBSERVATIONAL'
     *  ,' DATA  IS=? (IS <= 5)')
103   FORMAT(8X,'MAXIMUM LAG M = ?')
104   FORMAT(2X,'DIMENSIONAL INDEX OF OBSER'
```

```
     *  ,'VATIONS TAKEN TO BE'
     *  /3X,'DEPENDENT VARIABLE  KK = ?')
105     FORMAT(//,4X,'THE OPTIMAL S.M.L.R.'
     *  ,' MODEL IS:')
106     FORMAT(2X,4HX(K,,I1,2H)=,F6.3)
107     FORMAT(15X,2H+(,F6.3,6H)*X(K-,I1,
     *  1H,I1,1H))
        CLOSE(1)
999     STOP
        END
C*********************************************************
        SUBROUTINE SUBLSM(NO,MRE,K,A,ZK)
        DIMENSION A(MRE),X(1000,5),L(1000)
        COMMON /COM1/N,IS,X
        COMMON /COM2/M,KK,L
        LL=1
        NO=0
        DO 10 I=1,IS
        DO 10 J=1,M
        J1=K-J
        IF(J1.LE.0.OR.J1.GT.N)NO=NO-1
        A(LL)=X(J1,I)
10      LL=LL+1
        ZK=X(K,KK)
        RETURN
        END
C*********************************************************
        SUBROUTINE RECOV(N1,N2,MRE,S,E,SS,A)
        REAL S(MRE,MRE),E(MRE),SS(MRE),A(MRE)
        DIMENSION X(1000,5),L(1000)
        COMMON /COM1/N,IS,X
        COMMON /COM2/M,KK,L
        NO=0
        EZ=0.
        ZZ=0.
        MRE1=MRE-1
        DO 10 I=1,MRE
        SS(I)=0.
        E(I)=0.
        DO 10 J=I,MRE
10      S(I,J)=0.
        DO 11 K=N1,N2
```

```
      CALL SUBLSM(NO,MRE1,K,A,ZK)
      EZ=EZ+ZK
      ZZ=ZZ+ZK*ZK
      DO 21 I=1,MRE1
      AKI=A(I)
      S(I,MRE)=S(I,MRE)+AKI*ZK
      E(I)=E(I)+AKI
      SS(I)=SS(I)+AKI*AKI
      DO 22 J=I,MRE1
      AKJ=A(J)
22    S(I,J)=S(I,J)+AKI*AKJ
21    CONTINUE
11    CONTINUE
      S(MRE,MRE)=ZZ
      TN=FLOAT(N2-N1+NO+1)
      EZ=EZ/TN
      ZZ=SQRT(ZZ-TN*EZ*EZ)
      DO 14 I=1,MRE1
      T=E(I)/TN
      SS(I)=SQRT(SS(I)-TN*T*T)
14    E(I)=T
      SS(MRE)=ZZ
      E(MRE)=EZ
      DO 15 I=1,MRE
      T1=E(I)
      T2=SS(I)
      DO 15 J=I,MRE
      T=(S(I,J)-TN*T1*E(J))/(SS(J)*T2)
      IF(J.NE.I)S(J,I)=T
15    S(I,J)=T
      RETURN
      END
C**************************************************************
      SUBROUTINE SELECT(MS,S,XX,ES,F,FF,
     * BICO,II,ISW)
      DIMENSION S(MS,MS),XX(MS),ES(MS),
     * F(MS),FF(MS),L(1000),X(1000,5)
      COMMON /COM1/N,IS,X
      COMMON /COM2/M,RK,L
      ANM=FLOAT(N-M)
      MS1=MS-1
      YY=XX(MS)
```

```
      EY=ES(MS)
      DO 2 K=1,MS1
      L(K)=K
2     CALL ELIMIN(MS,MS,K,S,ISW)
      BICO=1.E8
      WRITE(*,102)
      DO 30 K1=1,MS
      K=MS-K1+1
      SO=(S(K,K)*YY**2)/ANM
      BIC=ALOG(SO)+2.*ALOG(ANM)*
     *     FLOAT(K-1)/ANM
      SO=SQRT(SO)
      KO=K-1
      T1=EY
      IF(KO.EQ.0)GOTO 10
      DO 50 I=1,KO
      LL=L(I)
      T=S(I,K)*YY/XX(LL)
      FF(I)=T
50    T1=T1-T*ES(LL)
10    FF(K)=T1
      WRITE(*,103)KO,BIC,SO
      IF(KO)51,51,52
51    WRITE(*,105)FF(K)
      GOTO 53
52    WRITE(*,104)(L(I),I=1,KO)
      WRITE(*,105)FF(K),(FF(I),I=1,KO)
53    IF(BIC.GT.BICO)GOTO 31
      SOO=SO
      BICO=BIC
      II=K-1
      II1=II+1
      F(II1)=FF(II1)
      IF(II.EQ.0)GOTO 31
      DO 1 I=1,II
1     F(I)=FF(I)
31    PO=1.E8
      IF(K.EQ.0)GOTO 30
      DO 3 I=1,KO
      P=S(I,K)**2/S(I,I)
      IF(PO.LT.P)GOTO 3
      PO=P
```

```
          IO=I
3         CONTINUE
          IF(K.EQ.1)GOTO 9
          DO 5 I=1,K
          T=S(I,IO)
          S(I,IO)=S(I,KO)
          S(I,KO)=S(I,K)
5         S(I,K)=T
          DO 7 I=1,K
          T=S(IO,I)
          S(IO,I)=S(KO,I)
          S(KO,I)=S(K,I)
7         S(K,I)=T
          IF(IO.EQ.K)GOTO 9
          LL=L(IO)
          L(IO)=L(KO)
          L(KO)=LL
          T=F(IO)
          F(IO)=F(KO)
          F(KO)=T
9         CALL [illegible]IN(MS,K,K,S,ISW)
          IF(ISW.NE.1)GOTO 62
30        CONTINUE
          WRITE(*,106)II,BICO,SOO
          IF(II)60,60,61
60        WRITE(*,108)F(II1)
          GOTO 62
61        WRITE(*,107)(L(I),I=1,II)
          WRITE(*,108)F(II1),(F(I),I=1,II)
102       FORMAT(2X,2HK,5X,3HBIC,8X,'R.E.S'
     *    ,3X,'SELECTED ORDER OF VARIABLES'
     *    ,3X,'PARAMETERS OF MODEL')
103       FORMAT(1X,I2,2(2X,E9.4))
104       FORMAT(25X,10I3)
105       FORMAT(57X,10F6.3)
106       FORMAT(/,3X,'OPTIMAL MODEL IS.',
     *    9X,'K=',I2,2X,'BIC=',E9.4,2X,
     *    'R.M.S=',E9.4)
107       FORMAT(3X,'SELECTED ORDER ',
     *    'VARIABLES',2X,10I3)
108       FORMAT(3X,'PARAMETERS OF MODEL'
     *    ,7X,10F6.3)
```

```
62        RETURN
          END
C*********************************************************
          SUBROUTINE ELIMIN(MM,M,K,S,ISW)
          REAL S(MM,MM)
          ISW=0
          IF(S(K,K).EQ.0.)GOTO 15
          S(K,K)=1./S(K,K)
          T1=S(K,K)
          DO 10 I=1,M
          IF(I.EQ.K)GOTO 10
          S(I,K)=-S(I,K)*T1
          S(K,I)=S(K,I)*T1
10        CONTINUE
          DO 12 I=1,M
          IF(I.EQ.K)GOTO 12
          DO 11 J=1,M
          IF(J.EQ.K)GOTO 11
          S(I,J)=S(I,J)+S(I,K)*S(K,J)/T1
11        CONTINUE
12        CONTINUE
          RETURN
15        ISW=0
          WRITE(*,102)ISW
102       FORMAT(3X,4HISW=,I1)
          RETURN
          END

C                   *******************
C                   *                 *
C                   *   27.   TREG    *
C                   *                 *
C                   *******************

          DIMENSION L(1000),X(1000,5),S(50,50),
     *    Y(50),D(50),ES(50),XX(50),A(50)
          COMMON /COM1/N,IS,X,NP1,ID,L
          OPEN(1,FILE='  ')
          WRITE(*,300)
          READ(*,*)N
```

```
      WRITE(*,302)
      READ(*,*)IS
      IF(IS-1)10,10,11
10    WRITE(*,303)
      READ(*,*)M
      GOTO 12
11    M=IS-1
12    WRITE(*,306)
      READ(*,*)ID
      WRITE(*,307)
      READ(*,*)KK1
      WRITE(*,308)
      READ(*,*)KKD
      WRITE(*,309)
      READ(*,*)SP1,SP2
      READ(1,*)((X(I,J),I=1,N),J=1,IS)
      CALL ORDER(KKD)
      CALL TREG(KKD,M,SP1,SP2,S,Y,A,AO,
     *     D,AIC,MO,ES,XX,ISW)
      CLOSE(1)
300   FORMAT(3X,'NUMBER OF DATA'
     *     ,' IN EACH DIMENSION   N=?')
309   FORMAT(3X,'SWITCHING POINTS   SP1=?',
     *     '   SP2=?')
302   FORMAT(3X,'DIMENSION OF OBSERVATI',
     *     'ONAL DATA   IS=?')
303   FORMAT(3X,' PRESETED MAXIMUM ORDER'
     *     ,' OF THRESHOLD'/,' AUTOREGRESSIVE ',
     *     'MODEL   M=?')
306   FORMAT(3X,' MAXIMUM LAG OF ',
     *     'THRESHOLD VARIABLE   ID=?')
307   FORMAT(3X,' THE DIMENSIONAL INDEX'
     *     ' OF OBSERVATIONS TAKEN',/,10X,
     *     'TO BE DEPENDENT VARIABLE   KK1 = ?')
308   FORMAT(3X,' THE DIMENSIONAL INDEX OF'
     *     ,/' THRESHOLD VARIABLE   KKD=?')
      STOP
      END
C*****************************************************
       SUBROUTINE SUBISM(NO,MRE,K,A,ZK)
       DIMENSION A(MRE),X(1000,5),L(1000)
       COMMON /COM1/N,IS,X,KK1,ID,L
```

```
      LK=L(K)
      LD=LK+ID
      IF(LD.GT.N)GOTO 22
      IF(IS-1)15,15,16
15    MRE1=MRE
      IF(LD-MRE.LE.0)GOTO 22
      GOTO 17
16    MRE1=MRE+1
17    DO 23 I=1,MRE1
      IF(IS-1)10,10,11
10    AO=X(LD-I,IS)
      I1=I
      GOTO 20
11    IF(I-KK1)12,23,13
12    I1=I
      GOTO 14
13    I1=I-1
14    AO=X(LD,I)
20    A(I1)=AO
23    CONTINUE
      ZK=X(LD,KK1)
      GOTO 99
22    NO=NO-1
      DO 18 I=1,MRE
18    A(I)=0.
      ZK=0.
99    RETURN
      END
C**********************************************************
       SUBROUTINE TREG(KKD,M,SP1,SP2,S,Y,A,
     * AO,D,AIC,MO,ES,XX,ISW)
      REAL S(M,M),Y(M),D(M),A(M),ES(M),XX(M)
      DIMENSION X(1000,5),L(1000)
      COMMON /COM1/N,IS,X,KK1,ID,L
      TN=FLOAT(N)
      NA=IFIX(TN*SP1)+1
      NB=IFIX(TN*SP2)+1
      NA1=NA
      IF(NA.NE.1)NA1=NA+1
      AICO=1E8
      IF(IS.GT.1)GOTO 11
      DO 1 MP=1,M
```

```
       CALL LSME(NA1,NB,MP,S,Y,ES,XX,A,
     * AO,D,SO,AIC,ISW)
       IF(ISW.EQ.1)GOTO 999
       IF(AICO.LE.AIC)GOTO 1
       AICO=AIC
       MO=MP
1      CONTINUE
       GOTO 12
11     MO=M
12     CALL LSME(NA1,NB,MO,S,Y,ES,XX,A,
     * AO,D,SO,AIC,ISW)
       IF(ISW.EQ.1)GOTO 999
       LNA=L(NA1)
       LNB=L(NB)
       TNA=X(LNA,KKD)
       TNB=X(LNB,KKD)
       WRITE(*,103)
       WRITE(*,102)SP1,SP2,TNA,TNB,AIC,
     * SO,MO,AO,(A(I),I=1,MO)
103    FORMAT(1X,'SWITCHING  SWITCHING',2X,
     * 'BIC',2X,'S.D.',1X,'T.R.-',4X,'T.R.-'
     * /,1X,'POINTS',5X,'POINTS',15X,'ORDER'
     * ,4X,'PARAME.'/,1X,'SP1 SP2',4X,'IN DATA')
102    FORMAT(1X,F3.1,1X,F3.1,4X,F4.0,1X,F4.0,
     * 1X,F4.1,1X,F5.1,2X,I2,4X,8F6.2)
999    RETURN
       END
C*********************************************
       SUBROUTINE ORDER(KKO)
       DIMENSION L(1000),X(1000,5)
       COMMON /COM1/N,IS,X,KK1,ID,L
       DO 22 K=1,N
22     L(K)=K
       N1=N-1
       DO 21 I=1,N1
       I1=L(I)
       XO=X(I1,KKO)
       DO 20 K=I,N
       J=L(K)
       X1=X(J,KKO)
       IF(XO.LT.X1)GOTO 20
       XO=X1
       KO=K
```

```
20        CONTINUE
          IF(KO.EQ.I)GOTO 21
          LO=L(KO)
          L(KO)=L(I)
          L(I)=LO
21        CONTINUE
          RETURN
          END

C                   ******************
C                   *                *
C                   *  28.   BILIM   *
C                   *                *
C                   ******************

          REAL X(1000,5),C(10),F(10),CC(5,5),
     *    S(45,45),SZ(45),SS(45),FF(45),A(45)
          COMMON /COM1/N,IS,X,IP,IQ,L,KK
          OPEN(1,FILE=' ')
          WRITE(*,101)
101       FORMAT(2X,'NUMBER OF DATA ',
     *    ' N=?  ')
          READ(*,*)N
          WRITE(*,102)
102       FORMAT(1X,'AR- AND MA- ORDER ',
     *    3X,'IP=?  IQ=?')
          READ(*,*)IP,IQ
          WRITE(*,103)
103       FORMAT(6X,'L=?    KK=?')
          READ(*,*)L,KK
          WRITE(*,104)
104       FORMAT(2X,'TOLERANCE  EPS=?')
          READ(*,*)EPS
          READ(1,*)(X(I,1),I=1,N)
          IS=2
          MAX=MAX0(IP,IQ,L,KK)
          MRE=IP+IQ+L*KK
          II=0
          WRITE(*,202)
41        II=II+1
          CALL RESID(MAX,FO,F,C,CC)
```

```
      CALL LSME(1,N,MRE,S,SZ,E,SS,FF,F0,
     *    A,S0,AIC,ISW)
      DAT=0.
      IF(IP.EQ.0)GOTO 31
      DO 30 I=1,IP
      ABSF=ABS(F(I)-FF(I))
      IF(ABSF.GT.DAT)DAT=ABSF
 30   F(I)=FF(I)
 31   IF(IQ.EQ.0)GOTO 32
      DO 33 I=1,IQ
      ABSF=ABS(C(I)-FF(I+IP))
      IF(ABSF.GT.DAT)DAT=ABSF
 33   C(I)=FF(I+IP)
 32   I1=0
      IF(L.EQ.0.OR.KK.EQ.0)GOTO 34
      DO 35 I=1,L
      DO 36 J=1,KK
      ABSF=ABS(CC(I,J)-FF(IP+IQ+I1+J))
      IF(ABSF.GT.DAT)DAT=ABSF
 36   CC(I,J)=FF(IP+IQ+I1+J)
 35   I1=I1+KK
 34   WRITE(*,201)II,S0,(F(I),I=1,IP),
     * (C(I),I=1,IQ),((CC(I,J),J=1,KK),I=1,L)
 201  FORMAT(2X,I2,4X,E8.3,10F8.3)
 202  FORMAT(1X,'ITERAT.     S.E.     F(I)'
     * 4X,  'CC(I,J)',/,3X,'No.')
      IF(DAT.GT.EPS)GOTO 41
      CLOSE(1)
      STOP
      END
C*******************************************************
      SUBROUTINE SUBLSM(NO,MRE,K,A,ZK)
      REAL X(1000,5),A(MRE)
      COMMON/COM1/N,IS,X,IP,IQ,L,KK
      ZK=X(K,1)
      IF(IP.EQ.0)GOTO 15
      DO 10 I=1,IP
      XI=0.
      IF(K-I.GT.0)XI=X(K-I,1)
 10   A(I)=XI
 15   IF(IQ.EQ.0)GOTO 16
      DO 11 I=1,IQ
      XI=0.
```

```
      IF(K-I.GT.0)XI=X(K-I,2)
11    A(IP+I)=XI
16    I1=0
      IF(L.EQ.0.OR.KK.EQ.0)GOTO 17
      DO 13 I=1,L
      XI1=0.
      IF(K-I.GT.0)XI1=X(K-I,1)
      DO 14 J=1,KK
      XI2=0.
      IF(K-J.GT.0)XI2=X(K-J,2)
14    A(IP+IQ+I1+J)=XI1*XI2
13    I1=I1+KK
17    RETURN
      END
C***************************************************************
      SUBROUTINE RESID(MAX,FO,F,C,CC)
      REAL F(MAX),C(MAX),CC(MAX,MAX),
     *  X(1000,5)
      COMMON /COM1/N,IS,X,IP,IQ,L,KK
      DO 20 IT=1,N
      ET=FO+X(IT,1)
      DO 10 I=1,MAX
      XI=0.
      EI=0.
      ITI=IT-I
      IF(ITI.LE.0)GOTO 18
      XI=X(ITI,1)
      EI=X(ITI,2)
18    IF(I.GT.L)GOTO 13
      DO 12 J=1,KK
      EJ=0.
      IF(IT-J.GT.0)EJ=X(IT-J,2)
12    ET=ET-CC(I,J)*XI*EJ
13    IF(I.GT.IP)GOTO 16
      ET=ET-F(I)*XI
16    IF(I.GT.IQ)GOTO 10
      ET=ET+EI*C(I)
10    CONTINUE
20    X(IT,2)=ET
      RETURN
      END
```

```
C                   ******************
C                   *                *
C                   *  29.   EXPOM   *
C                   *                *
C                   ******************

      REAL X(1000,5),F(50),E(50),SS(50),
     * S(50,50),A(50),SZ(50)
      COMMON /COM1/N,IS,X,IN,GAMA
      OPEN(1,FILE=' ')
      WRITE(*,101)
101   FORMAT(2X,'NUMBER OF DATA  N=?')
      READ(*,*)N
      WRITE(*,102)
102   FORMAT(2X,' IN=?  GAMA=?')
      READ(*,*)IN,GAMA
      READ(1,*)(X(I,1),I=1,N)
      IS=1
      MRE=2*IN
      CALL LSME(1,N,MRE,S,SZ,E,SS,F,FO,A,
     *   SO,AIC,ISW)
      WRITE(*,105)GAMA,SO
      WRITE(*,107)FO,(F(I),I=1,IN)
      WRITE(*,106)(F(I),I=IN+1,MRE)
105   FORMAT(2X,'GAMA=',3X,F4.2,'S.D.=',E9.3)
107   FORMAT(2X,'ALPH  : ',12F6.2)
106   FORMAT(2X,'BETA  : ',6X,11F6.2)
      CLOSE(1)
      STOP
      END
C*************************************************************
      SUBROUTINE SUBLSM(MRE,K,A,ZK)
      REAL A(MRE),X(1000,5)
      COMMON /COM1/N,IS,X,IN,GAMA
      ZK=X(K,1)
      DO 10 I=1,IN
      KI=K-I
      XK=0.
      IF(KI.GT.0)XK=X(KI,1)
      A(I)=XK
      A(IN+I)=EXP(-GAMA*XK*XK)*XK
```

```
10        CONTINUE
          RETURN
          END
C                 ******************
C                 *                *
C                 *  30.  ADAPAR   *
C                 *                *
C                 ******************

          REAL MEW,X(20),XX(5000),F(20,20),XL(20)
          OPEN(1,FILE=' ')
          WRITE(*,101)
101       FORMAT(2X,'NUMBER OF DATA----NN=?')
          READ(*,*)NN
          WRITE(*,102)
102       FORMAT(2X,'AUTOREGRESSIVE ORDER',
     *    '---N=?')
          READ(*,*)N
          WRITE(*,103)
103       FORMAT(2X,'MAX-STEP OF FORECAST---L=?')
          READ(*,*)L
          WRITE(*,104)
104       FORMAT(2X,'PARAMETER---MEW=?')
          READ(*,*)MEW
          READ(1,*) (XX(I),I=1,NN)
          WRITE(*,109)
109       FORMAT(3X,'OPTION FOR INITIAL ',
     *    'PARAMETERS'/5X,'KK1=? ( 0 = ALL'
     *    ,' ZERO  , 1 = NOT ZERO )')
          READ(*,*)KK1
          IF(KK1)14,14,15
14        DO 10 I=1,20
          DO 10 J=1,20
10        F(I,J)=0.
          GOTO 16
15        WRITE(*,107)
107       FORMAT(4X,'INITIAL VALUE OF '
     *    'AR-PARAMETERS---F(I,J), J=1,...,N'
     *    ,', I=1,...,L')
          READ(*,*)((F(I,J),J=1,N),I=1,L)
16        NL=N+L
```

```
      WRITE(*,108)
108   FORMAT(2X,'OPTION FOR STARTING
     *  POINT',/5X,' KK2=? ( 0= FROM
     *  NO.1 , 1 = FROM NO.NL+1)')
      READ(*,*)KK2
      IF(KK2)20,20,21
20    DO 23 I=1,NL
23    X(I)=0.
      NO=1
      GOTO 22
21    WRITE(*,119) NL
119   FORMAT(3X,'INITIAL VALUES OF X(I),
     *  I=1,...,',I2,':')
      DO 24 I1=1,NL
      I=NL-I1+1
      X(I)=XX(I1)
24    WRITE(*,*) X(I)
      NO=NL+1
22    DO 1 K=NO,NN
      X(1)=XX(K)
1     CALL ADAPAR(K,N,L,NL,X,XL,F,MEW)
      CLOSE(1)
      STOP
      END
C****************************************************************
      SUBROUTINE ADAPAR(K,N,L,NL,X,XL,F,MEW)
      REAL MEW,X(NL),F(L,N),XL(L)
      DO 10 J1=1,NL
      IF(J1-L)30,30,31
30    J=L-J1+1
      E=X(1)
      DO 20 I=1,N
      I1=I+J
20    E=E-F(J,I)*X(I1)
      E=2.0*MEW*E
      XLO=0.
      DO 21 I1=1,N
      I=N-I1+1
      I2=I+J
      YO=F(J,I)+E*X(I2)
      XLO=XLO+YO*X(I)
      IF(J-1)22,22,23
```

```
23        F(J,I)=F(J-1,I)
   )      GOTO 21
22        F(1,I)=YO
21        CONTINUE
          XL(J1)=XLO
31        JJ=NL-J1+1
          IF(JJ.EQ.1)GOTO 10
          X(JJ)=X(JJ-1)
10        CONTINUE
          WRITE(*,101)K,(F(1,I),I=1,N),X(1)
     *    ,(XL(I),I=1,L)
101       FORMAT(1X,I4,3X,F10.4,2X,F10.4,
     *    10F10.4)
          RETURN
          END

C>
```

参 考 资 料

（1） 周概容，概率论与数理统计.

(2) M.B. Priestley, Spectral Analysis and Time Series, London, 1981.

(3) M.B. Priestley, The Spectrum of a continuous process derived from a discrete process. Biometrika, 50, pp517-520.

（4） 安鸿志，陈兆国，杜金观，潘一民，时间序列的分析与应用，科学出版社，1983。

（5） 张尧庭，方开泰，多元统计分析引论，科学出版社，1982.

(6) S.M.Pandit, S.M.Wu, Time Series and System Analysis with Applications, John wiley and Sons, 1983.

(7) G.E.P. Box, G.M. Jenkins, Time Series Analysis Forecasting and Control, Rev.Ed.Holden-Day, 1976.

(8) An Hong-Zhi, Chen Zhao-Guo, E. J. Hannan, Autocorrelation, autoregression and autoregressive approximation, The Annals of Statistics, Vol.10, No.3, pp 926-936, 1982.

(9) S.A.Tretter, Introduction to Discrete-Time Signal Processing, John Wiley and Sons, 1976.

(10) R.K. Otnes, L. Enochson, Digital Time Series Analysis, Wiley-Interscience, 1972.

(11) N.C.Geckinli, D.Yavuz, Some novel windows

and a concise tutorial comparison of window families. IEEE Trans. ASSP, 26, pp 501-507,1978.

(12) N.C, Geckinli, D.Yavuz, Diserete Fourier Transformation and Its Applications to Power Spectra Estimation, Amsterdam,1983.

(13) 杜金观，潘一民，协方差可分离随机序列的新息与滤波，数学学报，第20卷，第1期，pp 16-27, 1977年3月。

(14) H.Akaike. AR Model fitting for control,Ann. Inst. Statist. Math. Vol.23, pp 163-180,1971.

(15) E.J. Hannan, Multiple Time Series, John Wiley and Sons, 1970.

(16) H.B.Mann,A. Wald, On the statistical treatment of linear stochastic difference equations. Econometrica, 11, pp 173-220, 1943,

(17) R. Shibata, Selection of the order of an autoregressive model by Akaike's information criterion. Biometrika, 63, pp 117-126, 1976.

(18) H. Akaike, Fitting autoregressive models for prediction, Ann. Inst.Statist. Math.,21, pp 243-247, 1969.

(19) H.Akaike, A Bayesian analysis of the minimum AIC procedure. Ann. Inst. Statist. Math., 30A, pp 9-14, 1978.

(20) H.Akaike, A Bayesian extension of the minimum AIC procedure of autoregressive model fitting. Biometrika, 66, pp 237-242, 1979.

(21) J.S.Bendat,A.G.Piersol,Random Data:Analysis and Measurement Procedures,Wiley-Interscience,1971.

(22) 顾岚，安鸿志，自回归模型的精细结构与统计分析，应用数学学报，第8卷，第四期，pp 433-445，1985年10月。

(23) H.Tong, K.S. Lim, Threshold autoregression, limit cycles and cyclical data, J. of Royal Statistical Society, B, 42, pp 245-292, 1980.

(24) 钟秋海，黎新华，张志方，一类非线性模型建模方法及其在铁路客流量预报上的应用，控制理论与应用，第2卷，第1期，pp 103—111，1985年1月。

(25) B. Widrow, Adaptive Filters I: Foundamentals, Rept. SEL-66-126 (TR6764-6), Stanford Electronics Laboratories, Stanford, CA, 1966.

(26) 杨位钦，机械颤动过程的适应性建模及其表征（英文），控制理论与应用，第1卷，第3期，pp 68-77，1984年7月。

(27) E.J. Hannan and J.Rissanen, The recursive estimation of mixed autoregressive-moving average, Biometrika, 69, pp 81-94, 1982.

